Journal of
CHRONIC
DISEASES

VOLUME 13 . NUMBER 5

MAY, 1961

ol. 13
1961

Symposium

DERMATOLOGICAL PROBLEMS

Special Editor
ERNEST W. SMITH, M.D.

INTRODUCTION

THE past decade of medical progress has seen a new approach to the basic mechanism of many cutaneous reactions. Clinical entities have been regrouped, in many instances, upon the firm foundation of similarity of pathogenesis rather than individualized upon the strength of minor morphologic variations.

The elucidation of biochemical mechanisms has been based to a large extent upon the application of histochemical techniques to cutaneous structures. An enormous and instructive literature has accumulated which bears upon the carbohydrate metabolism, energy transfer, and cellular localization of processes involved in epidermal growth and function. Histochemical characterization of connective tissue components has led to discrimination in dermal tissues found in normal and disease states.

Involvement of connective tissues also predominates in the major areas considered by the other contributors to this Symposium. The use of a myriad of terms for clinical variants of collagen-vascular, or connective-tissue, or rheumatoid diseases has long obscured a clear understanding of their interrelation. The revised and amplified concept of cutaneous manifestations of these disorders with a concise and logical orientation based upon current knowledge of pathogenesis should be helpful to everyone.

An area which, in the past, has been more obscure than the collagen-disease area involves light sensitivity. Of increasing clinical importance because of the increased use of photosensitizing substances in our culture and of basic physiologic interest because of the mechanisms of energy transport involved, reactions to various forms of light energy have been admirably classified in this Symposium.

Ernest W. Smith, M.D.
Baltimore, Md.

HISTOCHEMISTRY OF THE SKIN

RICHARD B. STOUGHTON, M.D.*
CLEVELAND, OHIO

From the Department of Dermatology, Western Reserve University

(Received for publication Nov. 22, 1960)

THE majority of histochemical investigations of skin have been conducted in the past ten years. The most complete review is a text by Montagna,[1] but much important work has been published since this excellent review.

In this paper discussion of the histochemistry of skin will be restricted to those techniques which attempt to show specific chemical entities in skin with specific methods which appear rational and are for the most part acceptable as specific chemical methods for identification. This, of course, eliminates all the literature dealing with such nonspecific staining methods as hematoxylin and eosin, orcein, Gomori trichrome, van Gieson, pH-sensitive dyes, etc.

There are still substantial limitations on the chemicals which can be specifically identified in skin by histochemical staining techniques. Among the agents which can be identified in skin are polysaccharides, fats, sulfhydryl and disulfide groups of proteins, free amino groups of proteins and lipoproteins, calcium, iron, ribonucleic acid, desoxyribonucleic acid, and a large group of oxidative and hydrolytic enzymes. The identification of enzymes by histochemical techniques is becoming one of the largest segments of the science of histochemistry.

The main advantage of histochemical techniques over strictly biochemical methods is that one is able to localize directly in the cell where the chemical in question resides. An example of the advantage of histochemical technique is the identification directly of the melanocyte as the cell in the epidermis which contains tyrosinase.[2] By exposing prepared human skin to tyrosine one can show that melanin is formed in the melanocytes and not in other cells of the epidermis. It is also possible with autoradiographs to show that melanocytes alone contain tyrosinase.[3] It must be realized that conventional biochemical techniques are desirable whenever strict localization is not part of the problem. The histochemist and his subject are hybrids of the genes of chemistry and anatomy. A deficiency in either results in a gross distortion of the product.

Original work in histochemistry was published by Unna,[4] a prolific dermatologist who contributed a great deal to the origin of this field. Lison[5] was an important later contributor. The field has been developed mostly in the past 15 years, with such contributors as Gomori,[6] Pearse,[7] Seligman and Lillie,[8] and others forming a nucleus for this growing branch of study.

*Director of Dermatology, Western Reserve University, Cleveland, Ohio.

Unfortunately, relatively little has been developed in the use of true histochemistry as a diagnostic procedure in evaluation of human disease. This is not surprising when one realizes the infancy of the field along with the fact that special fixation procedures are, so far, rather impractical for routine diagnosis. However, the field is one of growing importance in clinical pathologic study. A good example is the use of histochemical demonstration of tyrosinase as a measure of malignancy in melanocytic tumors of humans.[9] Also the PAS stain for polysaccharides has reached some degree of acceptance as a diagnostic procedure.[10] There are a number of others and in the near future there will be many more.

Fixation is a major technical problem of histochemistry, but we will not be able to discuss it in detail. In brief, formalin-fixed tissue is rarely useful for demonstration of enzymes because they are denatured too easily. Many techniques require such difficult and elaborate methods as frozen sections without fixation and frozen-dried preparations.

The use of electron microscopy in combination with specific chemical identification by "staining" should also be mentioned. This is a new field and at this time there is much investigation of the technique but little solid information.[11]

Interpretation of histochemical studies is sometimes difficult. Negative evidence, i.e., the absence of a material which should stain by a given technique, does not necessarily mean the substance is absent, but rather that it was not demonstrated. Use of different fixatives, particularly with enzyme identification, may result in completely different results. Localization is always open to question because intermediate products or even final products of a reaction may have migrated during the staining procedure, thus giving misleading ideas of localization. Any histochemical stain is extremely difficult to quantitate and for many reasons quantitation of chemical agents in tissue on the basis of histochemical techniques is open to serious question.

It seems easiest and perhaps most rational to discuss the histochemistry of *specific chemical agents* in skin rather than to discuss primarily *specific substructures of skin* and their histochemistry.

SULFHYDRYL AND DISULFIDE GROUPS

These groups are prominent and important in the physical structure of keratin, which occurs in hair, nails, and the entire outermost covering of the body. They are important cross-links in the parallel chains of the keratin molecules. If the disulfide groups are disrupted, keratin becomes a much more soluble p·otein and more susceptible to enzymes, acid, and alkali.[12]

An early stain for SH groups utilizing nitroprusside was quite specific but faded in seconds and so was of little practical use. Then an organic mercurial (Bennett's reagent) which is quite specific for sulfhydryl groups was introduced.[13] Nonspecific agents depending on the reducing activity of sulfhydryl groups were introduced but because of lack of specificity are of limited use. One of the best reagents for detecting free sulfhydryl groups is the DDD reagent.[14] To demonstrate disulfide groups, the tissue is exposed to disulfide reducing agents such as potassium cyanide or thioglycollate, and the free sulfhydryl groups thus exposed

are stained by the above methods for free sulfhydryl groups. Obviously the SH groups which are revealed are insoluble (or they would have disappeared from the tissue during fixation and staining) and are probably attached to insoluble keratin or keratin-like proteins. One can check the specific nature of the stain by previously blocking the SH groups with such agents as n-ethyl maleimide or iodoacetic acid.

All epidermal cells contain both sulfhydryl and disulfide groups. There is probably an intense accumulation of sulfhydryl groups in the transitional areas where keratin is being formed (keratogenous zone of hair follicles, supragranular layer area of epidermis). Intense accumulations of disulfide groups can be shown in hair and in epidermal keratin.[15]

POLYSACCHARIDES

It is best to consider two general groups of polysaccharides and deal with them separately in skin:

A. The ground substance and its contents, which consist primarily of polymers of glucuronic acid and acetyl glucosamine (hyaluronic acid); and glucuronic acid, acetyl galactosamine, and sulfate (chondroitin sulfuric acid).

B. Glycogen and its various forms.

A. The Ground Substance.—There are a number of somewhat specific stains for ground substance (Hale, Bauer, PAS, toluidine blue, thionin, alcian blue) but the most widely used and probably most generally dependable is the PAS (periodic acid–Schiff) stain.[16,17] The specificity resides in the fact that periodic acid will oxidize polysaccharides by disrupting adjacent carbon atoms (C_2 and C_3) on which OH or NH_2 groups are present.[18,19] Aldehydes are formed by this reaction and these aldehydes are subsequently detected by Schiff's reagent. Chondroitin sulfuric acid and fully sulfated heparin do not react with this technique but monosulfated heparin will.[20]

The PAS stain is also an effective method by which to demonstrate bacteria and fungi in tissue.[21]

Using amylase to treat control specimens, one can differentiate glycogen from other PAS-positive material. We will first consider amylase-resistant PAS-positive material (presumably polysaccharides other than glycogen) in skin. Throughout the fibers of the corium there are deposits of polysaccharide material which is resistant to hyaluronidase but is apparently hydrolyzed by collagenase.[10,22] Similar PAS material is found more densely deposited at the junction of epithelial cells and mesodermal derivatives. Thus, surrounding the sweat glands, follicles, and sebaceous glands and at the dermal-epidermal junction, there are dense deposits of hyaluronidase-resistant, collagenase-sensitive material. Secretory cells of sweat glands may contain nonglycogen PAS-positive material which is extruded into the sweat ducts.[23,23a] Definite and rather characteristic alterations of the PAS-positive material of the dermal epidermal junction may be seen in lupus erythematosus.[10]

A skin lesion associated with hyperthyroidism is localized myxedema. The PAS-positive material of the corium in this tissue may not be grossly abnormal

even though there is a major increase in hyaluronic acid and chondroitin sulfuric acid. Treatment of a formalin-fixed section from localized myxedema with hyaluronidase will not alter the PAS staining, but it will completely abolish the pronounced metachromasia seen with such dyes as toluidine blue or thionin.[24]

Mast cells and young fibroblasts contain prominent granules which are amylase resistant and PAS positive. The material in the mast cells is thought to be heparin. Also, one worker has supported the idea that hyaluronic acid is contained in mast cells.[25] The PAS-positive material in young fibroblasts is probably some precursor of the polysaccharides of ground substance. The PAS material in fibroblasts is resistant to hyaluronidase.

The vessel walls contain large amounts of PAS-positive material which is amylase resistant. The PAS stain is an excellent one to outline the vessel distribution in sections of skin.

The fact that fungi contain PAS-positive material has been utilized to identify the organisms in fungus diseases of skin as well as other organs.[21] A galactolipid has been identified in lipoid proteinosis, a rare skin disease,[26] as well as in the lesions of Gaucher's disease.

B. Glycogen.—Amylase-sensitive, PAS-positive material in human skin is usually assumed to be glycogen. Many investigations of glycogen in normal and diseased skin have been reported.[27-29] In general, the more active the epithelial proliferative process, the more likely is there to be glycogen in the epidermal cells. The presence of glycogen has been correlated with mitosis of the cell in the experiments by Lobitz.[30] Glycogen is seen in the resting sweat gland but disappears after activity of the gland.[31] Sweat contains significant amounts of lactic acid. All of this suggests that energy for sweat secretion comes, at least in part, from glycogen utilization.[15]

Glycogen is also seen regularly in sebaceous gland epithelium and follicular epithelium which are very active proliferative areas much of the time.

RIBONUCLEIC ACID (RNA) AND DESOXYRIBONUCLEIC ACID (DNA)

There are a number of ways to demonstrate DNA and RNA. Most simple is the Feulgen method for DNA. Desoxyribose is changed to an aldehyde by mild exposure to hydrochloric acid at 50° C. This aldehyde is developed by the Schiff reagent to reveal the site of DNA. This can be checked by treating a control slide with desoxyribonuclease which eliminates the Feulgen-positive material. Similar specific enzyme digestion of RNA can be accomplished with ribonuclease and the RNA can be specifically stained with pyronin.

These substances can be demonstrated in human skin cells as in cells of any other organ. There are no special differences from other cells that can be detected by histochemical techniques. In a study of certain pigmented skin tumors, it was thought that there was more Feulgen material (DNA) in basal cell carcinoma and malignant melanoma than in other pigmented tumors. Also clear cell nuclei of lentigo and junction nevus were more strongly Feulgen positive than other cells in the adjacent epidermis.[32]

CALCIUM

Normal human skin contains significant amounts of calcium and, according to interpretations of microincineration studies, the epithelial cells contain rather large amounts.[33] However, the conventional histochemical stains for calcium fail to show this mineral in normal skin. In abnormal skin, such as that of scleroderma with calcium deposition, osteoma cutis, pseudoxanthoma elasticum, calcifying epithelioma, and others, calcium can often be easily detected with the von Kossa stain. This method actually depends on the displacement of calcium from its phosphate component by silver ions because of the greater insolubility of silver phosphate. The silver ions are then reduced to give a black metallic silver pigment at the site of calcium. For practical purposes this is a reliable stain for calcium but obviously not as specific in nature as one would like a true histochemical procedure to be.

IRON AND MELANIN

Of practical value in determining the nature of brown pigment in the skin are stains for iron and melanin. The two pigments may appear identical when routine staining methods are used. The old standard ferrocyanide method for iron is still one of the most reliable. Melanin may be distinguished by using ionic silver which is reduced to insoluble metallic silver by melanin granules. Thus iron stains blue and melanin black by these respective methods. In normal skin the ferrocyanide stain for iron is negative. However, melanin is usually present in normal skin in the basal cells and melanocytes as well as in a scattered but slight amount throughout the other epidermal cells and even the keratin layer. A positive iron stain usually indicates a previous escape of red blood cells from the vascular spaces. Melanin granules in an atypical location may indicate a pigmented nevus, previous inflammatory change, melanoma, and a number of disorders of melanin formation and distribution.

FREE AMINO GROUPS

Potter has recently reviewed the methods for obtaining free amino groups in tissue.[34] Using the ninhydrin-Schiff method,[35] he showed a rather general staining of epithelial cells and collagen bundles of human skin. This method is thought to distinguish the free amino groups of protein and lipoprotein in skin.

THE DEMONSTRATION OF ENZYMES

We will next consider this most exciting and rapidly advancing area of histochemistry. In general, and with few exceptions, the method involves the use of a water-soluble substrate which contains an ester similar, if not identical, to the physiologic substrate of the enzyme to be measured. One of the split products has to be insoluble or must immediately combine with an agent in the medium which renders it highly insoluble. This precipitated reaction product may be colorful in itself or may require further development to a colored product which can be seen microscopically. These prerequisites obviously frustrate the

histochemical detection of many enzymes but it is one of the more exciting fields of biologic science in which substrates are invented to fit the requirements for histochemical identification of a given enzyme. A good example is the detection of alkaline and acid phosphatase.[36] Alkaline and acid phosphatase have been shown biochemically to be separate enzymes,[37] and histochemically the localizations are quite different for alkaline and acid phosphatase. Gomori[38] first demonstrated alkaline phosphatase by histochemical technique. The early work involved a substrate of glycerol phosphate. Tissue phosphatase split this ester in an alkaline medium and released phosphate ions. Calcium present in the medium for the purpose immediately precipitated the phosphate in situ. This was a colorless product and not insoluble enough, so the tissue was treated with cobalt ions which replaced calcium to give a more insoluble cobalt phosphate. Then the tissue was treated with ammonium sulfide to form the more insoluble cobalt sulfide which is also a black color and easily seen with the light microscope. So the original enzyme is finally detected as an insoluble cobalt sulfide, neither ion of which was present in the original incubation medium.

A more recent method for demonstration of phosphatase is worth mentioning because this basic method has been used to demonstrate many hydrolytic enzymes.[6] Naphthol phosphate is used as the substrate. Included in the medium is a diazonium salt which is fairly stable and soluble. When the phosphate ion is split away from the naphthol part of the molecule, the naphthol combines readily with the diazonium salt to form an insoluble colored precipitate which indicates the site of phosphatase activity. Naphthol esters of sulfuric acid, glucuronic acid, galactopyranoside (b-d-galactosidase),[39] acetate, amino acids, and other biologic units have been used in detecting the specific hydrolytic enzyme involved.

Fixation of tissue is one of the main problems in demonstrating enzymes. Most enzymes will be destroyed by formalin fixation. Many enzymes will partially or almost completely survive alcohol or acetone fixation. Almost all enzymes will survive freezing and sectioning and this method is preferred whenever possible. Again, one must remember that inability to demonstrate an enzyme does not mean that the enzyme is not present. Ultracentrifugation techniques now allow biochemical analysis of component parts of the cells, but microscopic histochemistry is still superior in many respects in localizing enzymes within the cell and in special cells of various tissues.

It is not within the scope of this paper to review the detailed histochemical work in the field of cutaneous enzymes. I will try to cover a few interesting and significant conclusions that have been drawn from these numerous studies. An incomplete survey of the literature in the past eight years uncovered over 600 articles in this area alone.

A. Alkaline Phosphatase.—This is a widespread enzyme in tissues and also in skin. It is impossible to review all the reports on the presence of this enzyme in human skin. It is found in capillaries and has been used to outline vessels around sebaceous glands and hair follicles, eccrine sweat glands and ducts.[40] It is also found in most epithelial cells. Studies of wound healing showed that when keratinization was suppressed there was little acid or alkaline phosphatase

demonstrable.[41] It is well to remember that phosphatases at alkaline pH have been demonstrated with such substrates as 5-nucleotidase, glycerophosphate, and glucose-6-phosphate.[42]

B. Acid Phosphatase.—Since the original method was published by Gomori[43] there have been many conflicting reports on localization of this enzyme. Moretti and Mescon[44] have provided a good summary of the problem and report that acid phosphatase is concentrated in areas where keratin is being actively formed. They point out that this is also the area where SH groups are concentrated. It is known, however, that just above the granular layer there is a concentration of b-glucuronidase and some esterases.

Acid phosphatase is also present in sebaceous glands,[45] eccrine sweat glands,[46] and squamous cell carcinoma.[47]

C. Esterases.—These enzymes are an extremely broad group which catalyze the hydrolysis of esters of carboxylic acids. More specifically they will hydrolyze esters of N-free alcohols and phenols and esters of choline. This rather arbitrary simplification is really unfair because of the large number of esterases known to exist, their overlapping in regard to substrate specificity and inhibitors, and the endless subclassifications that have been proposed.[48] The article by Gomori[48] is worth reading for orientation regarding diversity of esterases in human tissues. A good review of the specificity of cholinesterases is presented by Chessick.[49] Suffice it to say that there is serious doubt about the histochemical differentiation between "true" and "pseudo" cholinesterase. However, this large group of eserine-sensitive, choline ester-splitting enzymes maintain some group identity although there is great variation from tissue to tissue and species to species.

In consonance with pharmacologic indications that sweat glands are controlled by acetylcholine liberated via sympathetic nerves, cholinesterase has been identified in great concentrations around the sweat glands.[50] Hurley and associates[50] found no true cholinesterase around sebaceous glands, follicles, epidermis, or apocrine glands. Aavik[51] claimed to find true or specific cholinesterase around apocrine glands and in pilomotor muscles of skin.

Montagna and Ellis[52] have shown cholinesterase in nerve plexi at the level of sebaceous glands and eccrine glands and infundibulum of the follicle.

Since the discovery of reactivation of cholinesterase by oximes,[53] a histochemical method showing this reactivation has been reported for the cholinesterase of the end plates.[54]

Another group of esterases includes those that hydrolyze short chain fatty acid esters (simple esterases) and long chain fatty acid esters (lipases). Findlay[55] studied the simple esterases of human skin both biochemically and histochemically, using the same substrates for both procedures. He showed that some esterase activity was demonstrable biochemically in human skin for naphthol esters of saturated fatty acids having from 2 to 10 carbon atoms. The peak activity seemed to be with esterase for the 4-5 carbon chains. Histochemically, he used acetic and butyric esters of naphthol as well as acetate and chloracetate esters of the anilide of 2-hydroxy-3-naphthoic acid (naphthol AS). Esterases were present in virtually every cell type found in skin. They were very concentrated in surface keratin, inner root sheath cuticle, and sweat duct cuticle. The concentration of esterases

on the surface may be related to the intense metabolic activity of surface lipids.[15]

Attempts to show lipases in human skin have been mainly disappointing. Steigleder,[56] using naphthol acetate, has demonstrated this esterase in fungi in skin and in surface wipings of skin. Material excreted via the pilosebaceous apparatus has strong esterase activity. Wells[57] has shown intense simple esterase activity in histiocytes containing Leishmania bodies in cutaneous leishmaniasis.

D. Phosphamidase.—Gomori[58] reported a peculiar enzyme in human skin which catalyzed the hydrolysis of a phosphorous-nitrogen linkage. He labeled this enzyme phosphamidase. It is found in malignant epidermal cells and not in normal quiescent epidermal cells. However, it may be found in nonmalignant proliferating epidermal cells so that it is of little specificity in detecting malignant cells. Meyer and Weinmann[59] found that phosphamidase was associated with keratin formation, being present in the granular layer and on out into the keratin.

E. Aminopeptidase.—Burstone and Folk[60] describe a method for determining leucine aminopeptidase. A leucine ester of naphthol is utilized as a substrate. Braun-Falco[61,62] has demonstrated this enzyme in normal skin, cantharidin blisters, and a number of dermatoses. It is difficult to quantitate this histochemical method and it has been shown that, although there is an increased concentration of this enzyme around malignancies, there is also increased activity around some benign lesions associated with fibroblast proliferation.[63]

F. Monamine Oxidase.—Shelley and co-workers[64] demonstrated monamine oxidase in epidermis, sebaceous gland, eccrine and apocrine glands. The role of monamine oxidase in metabolism of epinephrine, norepinephrine, and 5-hydroxytryptamine is under investigation.[65] This is of interest in regard to recent findings of Burch[66] which indicate that epinephrine-like compounds are present in human skin.

Recent work[67] indicates that o-methyl transferase rather than monamine oxidase is responsible for inactivation of epinephrine and norepinephrine.

G. Phosphorylase and Glucosidase.—These enzymes play a dominant role in the interconversion of glucose and glycogen. Phosphorylase alone can form 1,4 linkages between glucose moieties to produce straight chain polymers. Amylo 1,6 glucosidase forms 1,6 linkages for the development of branching chains. Together these enzymes synthesize the large branched polymer, glycogen.

These enzymes can be distinguished because of the different end products they produce, which stain differently with iodine.[68,69] These enzymes can also be differentiated on the basis of specific inhibitor substances. Their identification has been accomplished in human skin.[70,71] They have been found in (a) epidermis, (b) growing hair follicles (resting hair follicles contain only phosphorylase), (c) secretory and excretory sweat structures, which are rich in phosphorylase, and (d) sebaceous glands, which contain both enzymes.

Goltz[70] has recently shown histochemically that both succinic dehydrogenase and phosphorylase activity are present in the epidermis, and discusses the relationship of these enzymes to oxidative and anaerobic metabolism of carbohydrates in the epidermis. He found no qualitative differences in carbohydrate metabolism in normal skin and in cutaneous neoplasms.

Recently b-d-galactosidase has been identified histochemically in epithelial cells of the rat.[72] The substrate used was 6 bromo-b-d-galactopyranoside.

H. Succinic Dehydrogenase.—Dehydrogenase enzymes are widespread in tissues. Succinic dehydrogenase, an important enzyme in oxidative metabolism of carbohydrates, has been demonstrated in skin. A good review, as well as presentation of the author's own work, is to be found in an article by Montagna and Formisano.[73] They were able to demonstrate the enzyme rather easily in most epithelial cells. The activity seems more prominent in the germinative areas of the epidermis. Other techniques have demonstrated the presence in skin of most of the enzymes concerned with carbohydrate metabolism.[74,75]

REFERENCES

1. Montagna, W.: The Structure and Function of Skin, New York, 1956, Academic Press, Inc.
2. Fitzpatrick, T. B., Becker, S., Jr., Lerner, A., and Montgomery, H.: Tyrosinase in Human Skin: Demonstration of Its Presence and Its Role in Human Melanin Formation, Science **112**:223, 1950.
3. Fitzpatrick, T. B., and Kukita, A.: A Histochemical Autoradiographic Method for Demonstration of Tyrosinase in Human Melanocytes, Nevi, and Malignant Melanoma, J. Invest. Dermat. **26**:173, 1956.
4. Unna, P.: Histopathology of the Diseases of the Skin, New York, 1896, The Macmillan Company.
5. Lison, L.: Histochimie animale, Paris, 1936, Gautier-Villars.
6. Gomori, G.: Microscopic Histochemistry, Chicago, 1952, University of Chicago Press.
7. Pearse, A. G. E.: Histochemistry, Theoretical and Applied, Boston, 1953, Little, Brown & Company.
8. Lillie, R. D.: Histopathologic Technique and Practical Histochemistry, New York, 1954, McGraw-Hill Book Company, Inc.
9. Gordon, M., editor: Pigment Cell Biology, New York, 1959, Academic Press, Inc.
10. Stoughton, R. B., and Wells, G. C.: A Histochemical Study of Polysaccharides in Normal and Diseased Skin, J. Invest. Dermat. **14**:37, 1950.
11. Barnett, R. J., and Palade, G. E.: Application of Histochemistry to Electron Microscopy, J. Histochem. & Cytochem. **6**:1, 1958.
12. Lorincz, A. L., and Stoughton, R. B.: Specific Metabolic Processes in Human Skin, Physiol. Rev. **38**:481, 1958.
13. Mescon, H., and Flesch, P.: Modification of Bennett's Method for the Histochemical Demonstration of Free Sulfhydryl Groups in Skin, J. Invest. Dermat. **18**:261, 1952.
14. Barnett, R. J., and Seligman, A. M.: Histochemical Demonstration of Protein-Bound Sulfhydryl Groups, Science **116**:323, 1952.
15. Rothman, S.: Physiology and Biochemistry of the Skin, Chicago, 1954, University of Chicago Press.
16. McManus, J. F. A.: Histochemical Demonstration of Mucin After Periodic Acid, Nature **158**:202, 1946.
17. Hotchkiss, R. D.: Microchemical Reaction Resulting in the Staining of Polysaccharide Structures in Fixed Tissue Preparations, Arch. Biochem. **16**:131, 1948.
18. Nicolet, B. H., and Shinn, L. A.: The Action of Periodic Acid on Amino Alcohols, J. Am. Chem. Soc. **61**:1615, 1939.
19. Jackson, E. L., and Hudson, C. S.: The Structure of the Products of the Periodic Acid Oxidation of Starch and Cellulose, J. Am. Chem. Soc. **60**:989, 1938.
20. Jorpes, J. E., Werner, B., and Aberg, B.: The Fuchsin-Sulfurous Acid Test After Periodate Oxidation of Heparin and Allied Polysaccharides, J. Biol. Chem. **176**:272, 1948.
21. Kligman, A. M., and Mescon, H.: The Periodic Acid Schiff Stain for the Demonstration of Fungi in Animal Tissue, J. Bact. **60**:415, 1950.
22. Stoughton, R. B., and Lorincz, A. L.: The Action of Collagenase on Skin and the Anticollagenase Factor in Human Serum, J. Invest. Dermat. **16**:43, 1951.
23. Dobson, R. L., Formisano, V., Lobitz, W., and Brophy, D.: Some Histochemical Observations on the Human Eccrine Sweat Glands, J. Invest. Dermat. **31**:147, 1958.

23a. Hambrick, G. W.: Periodic Acid–Schiff Positive Material Accumulating Within the Lumen of Eccrine Sweat Glands, J. Invest. Dermat. **29**:213, 1957.

24. Stoughton, R. B.: Personal observations.
25. Asboe Hausen, G.: A Survey of the Normal and Pathologic Occurrence of Mucinous Substances and Mast Cells in the Dermal Connective Tissue, Acta dermat.-venereol. **30**:338, 1950.

26. Wood, M. G., Urbach, F., and Beerman, H.: Histochemical Study of a Case of Lipoid Proteinosis, J. Invest. Dermat. **26**:263, 1956.
27. Steiner, K.: Mucoid Substances and Cutaneous Connective Tissue in Dermatoses. I, II, and III. J. Invest. Dermat. **28**:387, 1957.
28. Goltz, R. W., Fusaro, R. M., and Jarvis, J.: Observations on Glycogen in Epithelial Tumors, J. Invest. Dermat. **31**:331, 1958.
29. Washburn, W. W.: Comparative Histochemical Observations on Wound Healing in Adult Rats and Cultured Adult Human Epithelium. I. Methods and Glycogen Distribution, J. Invest. Dermat. **23**:97, 1954.
30. Lobitz, W., and Holyoke, J.: The Histochemical Response of the Human Epidermis to Controlled Injury. Glycogen, J. Invest. Dermat. **22**:189, 1954.
31. Shelley, W., and Mescon, H.: Histochemical Demonstration of Secretory Activity in Human Eccrine Sweat Glands, J. Invest. Dermat. **18**:289, 1952.
32. Cawley, E. P., McManus, J. F. A., and Wheeler, C. E.: Observations on the Desoxyribonucleic Acid Component of Pigmented Skin Tumors, J. Invest. Dermat. **27**:9, 1956.
33. Lansing, A. I., and Opdyke, D.: Histological and Histochemical Studies of the Nipples of Estrogen-Treated Guinea Pigs With Special Reference to Keratohyalin Granules, Anat. Rec. **107**:379, 1950.
34. Potter, B.: Histochemical Demonstration of Free Amino Groups in Cutaneous Proteins, J. Invest. Dermat. **33**:245, 1959.
35. Yasuma, A., and Ichikawa, T.: Ninhydrin-Schiff and Alloxan-Schiff Staining, J. Lab. & Clin. Med. **41**:296, 1953.
36. Gomori, G.: Further Studies on the Histochemical Specificity of Phosphatases, Proc. Soc. Exper. Biol. & Med. **72**:449, 1949.
37. Perlman, G. E., and Ferry, R. N.: A Note on the Separation of Kidney Phosphatases, J. Biol. Chem. **142**:153, 1942.
38. Gomori, G.: Histochemistry of Enzymes, J. Mt. Sinai Hosp. **19**:446, 1952.
39. Cohen, R. B., Tsou, K., Rutenberg, S., and Seligman, A.: The Colorimetric Estimation and Histochemical Demonstration of B-D-Galactosidase, J. Biol. Chem. **239**:239, 1952.
40. Ellis, R. A., Montagna, W., and Fanger, H.: Histology and Cytochemistry of Human Skin. XIV. The Blood Supply of the Cutaneous Glands, J. Invest. Dermat. **30**:137, 1958.
41. Washburn, W. W.: Comparative Histochemical Observations on Wound Healing in Adult Rats and Cultured Adult Human Epithelium. III. Alkaline and Acid Phosphatase, J. Invest. Dermat. **24**:537, 1955.
42. Wachstein, M., and Meisel, E.: Substrate Specific Phosphatases at pH 7.2 in Biliary Obstruction and Liver Cell Damage, A.M.A. Arch. Path. **65**:449, 1958.
43. Gomori, G.: Distribution of Acid Phosphatases in the Tissues Under Normal and Pathologic Conditions, Arch. Path. **32**:189, 1941.
44. Moretti, G., and Mescon, H.: Histochemical Distribution of Acid Phosphatases in Normal Human Skin, J. Invest. Dermat. **26**:347, 1956.
45. Montagna, W., and Hamilton, J. S.: Histochemical Analysis of the Sebaceous Glands of the Hamster. Fed. Proc. **6**:166, 1947.
46. Shelley, W., and Mescon, H.: Histochemical Demonstration of Secretory Activity in Human Eccrine Sweat Glands, J. Invest. Dermat. **18**:289, 1952.
47. Lemon, H. M., and Wisseman, C. L.: Acid Phosphomonoesterase Activity of Human Neoplastic Tissue, Science **109**:233, 1949.
48. Gomori, G.: Human Esterases, J. Lab. & Clin. Med. **42**:445, 1953.
49. Chessick, R.: The Histochemical Specificity of Cholinesterases, J. Histochem. & Cytochem. **2**:258, 1954.
50. Hurley, H. J., Shelley, W. B., and Koelle, G. B.: The Distribution of Cholinesterases in Human Skin, With Special Reference to Eccrine and Apocrine Sweat Glands, J. Invest. Dermat. **21**:139, 1953.
51. Aavik, O.: Cholinesterases in Human Skin, J. Invest. Dermat. **24**:103, 1955.
52. Montagna, W., and Ellis, R.: Histology and Cytochemistry of Human Skin. XII. Cholinesterases in the Hair Follicles of the Scalp, J. Invest. Dermat. **29**:151, 1957.
53. Childs, A., Davies, D. R., Green, A. L., and Rutland, J. P.: Reactivation by Oximes and Hydroxamic Acids of Cholinesterase Inhibited by Organophosphorous Compounds, Brit. J. Pharmacol. **10**:462, 1955.
54. Bergner, A. O., and O'Neill, J.: A Modification of the Koelle Technique for Use With Oximes, J. Histochem. & Cytochem. **6**:72, 1958.
55. Findlay, G.: The Simple Esterases of Human Skin, Brit. J. Dermat. **67**:83, 1955.
56. Steigleder, G. K.: Histochemistry of Plantar Hyperkeratoses, J. Invest. Dermat. **31**:29, 1958.
57. Wells, G. C.: Hydrolysing Enzymes in Macrophages, J. Invest. Dermat. **31**:83, 1958.
58. Gomori, G.: Histochemical Demonstration in Sites of Phosphamidase Activity, Proc. Soc. Exper. Biol. & Med. **69**:407, 1948.
59. Meyer, J., and Weinmann, J. P.: Occurrence of Phosphamidase Activity in Keratinizing Epithelia, J. Invest. Dermat. **29**:393, 1957.

Vol. 13
1961

60. Burstone, M. S., and Folk, J. E.: Histochemical Demonstration of Aminopeptidase, J. Histochem. & Cytochem. **4**:217, 1956.
61. Braun-Falco, O., and Salfeld, K.: Über das Verhalten der leucin-amino peptidase-Aktivitat im Blutserum und blasenin Halt, Arch. klin. u. exper. Dermat. **205**:103, 1957.
62. Braun-Falco, O.: Histochemische aminopeptidase-durt-sellung in normaler Haut, bei Psoriasis, Dermatitis, Basaliom, spinozellulärem Karzinom und Molluscum sebaceum, Dermat. Wchnschr. **134**:104, 1956.
63. Monis, B., Nachlas, M. M., and Seligman, A. M.: Histochemical Observations on Leucine Aminopeptidase Activity in Diseased Tissues, J. Histochem. & Cytochem. **6**:395, 1958.
64. Shelley, W. B., Cohen, S. B., and Koelle, G. B.: Histochemical Demonstration of Monamine Oxidase in Human Skin, J. Invest. Dermat. **24**:561, 1955.
65. Sjoerdsma, A.: Alterations in the Pattern of Amine Excretion in Man Produced by a Monamine Oxidase Inhibitor, Science **130**:225, 1959.
66. Burch, G. E., and Phillips, J. H.: Chromaffin Reacting Cells in Human Digital Skin, Circulation Res. **6**:416, 1958.
67. Axelrod, J., and Laroche, M. J.: Inhibitor of O-methylation of Epinephrine and Norepinephrine in Vitro and in Vivo, Science **130**:800, 1958.
68. Takeuchi, T.: Histochemical Demonstration of Phosphorylase, J. Histochem. & Cytochem. **4**:84, 1956.
69. Takeuchi, T.: Histochemical Demonstration of Branching Enzyme (Amylo-1,4→ 1,6 Transglycosidase) in Animal Tissues, J. Histochem. & Cytochem. **6**:208, 1958.
70. Goltz, R. W., Fusaro, R. M., Blazejovsky, R., and Jarvis, J.: Pathways of Carbohydrate Metabolism in Normal and Neoplastic Cutaneous Ectoderm. A Histochemical Study, J. Invest. Dermat. (in press).
71. Ellis, R. A., and Montagna, W.: Histology and Cytochemistry of Human Skin. XI. Sites of Phosphorylase and Amylo-1,6 Glucosidase Activity, J. Histochem & Cytochem. **6**:201, 1958.
72. Rutenberg, A. M., Rutenberg, S. H., Monis, B., Teaque, R., and Seligman, A. M.: Histochemical Demonstration of B-D Galactosindase in the Rat, J. Histochem. & Cytochem. **6**:123, 1958.
73. Montagna, W., and Formisano, V. R.: Histology and Cytochemistry of Human Skin. VII. The Distribution of Succinic Dehydrogenase, Anat. Rec. **122**:1, 1955.
74. Pomerantz, S.: Personal communication.
75. Cruickshank, C. N. D., Hershey, F. B., and Lewis, C.: Isocitric Dehydrogenase Activity of Human Epidermis, J. Invest. Dermat. **30**:33, 1958.

SUNLIGHT AND THE SKIN

JOHN M. KNOX, M.D.
HOUSTON, TEXAS

From the Department of Dermatology, Baylor University College of Medicine

(Received for publication Nov. 22, 1960)

THE effect of sunlight on plants and animals has been a fascinating subject since ancient times. In some primitive cultures the sun was worshipped as a god, and down through the ages there has been more emphasis on the advantages than on the harm of sunlight. Dermatologists have led the way in studying the effects, both good and bad, of sunlight upon the skin. It is significant that in the Southwest dermatologists have become increasingly aware of the damage that can develop from prolonged exposure to sunlight.

The present trend of our American way of life is toward more time and money for recreation, and much of this additional relaxation time is being spent in outdoor activities. In the past, only farmers, sailors, and other outdoor workers were exposed to large amounts of sunlight over prolonged periods of time. With life expectancy continually being prolonged, the results of chronic irradiation are assuming more importance. Among the undesirable responses with which the physician must be familiar are sunburn, photosensitivity reactions, aging, and carcinogenesis. Physicians should be able to advise their patients in regard to the possible ill effects of sunshine and recommend appropriate means of protecting the skin from damage.

The term "sunlight" includes visible light, ultraviolet light, and infrared light. The majority of cutaneous changes produced by sunlight are caused by ultraviolet, for it is these rays that cause the greatest biologic effect on the skin. However, photoallergic, phototoxic, and photodynamic reactions may be produced by longer wave lengths. Sunlight is composed of electromagnetic waves that vary in length, and at the earth's surface these wave lengths range from 2,900 to 18,500 Å. The visible range lies between 3,900 and 7,400 Å, with ultraviolet the shorter, and infrared the longer, wave lengths.[1] Although not classified as a form of ionizing irradiation, ultraviolet photons can produce some ionization. Photons of ultraviolet energy are absorbed by the skin and there is little or no penetration beyond a few millimeters.[2]

The purpose of this paper is to provide the reader with fundamental information related to sunlight and the skin. Most of the significant diseases in which sunlight is a causative or contributing factor will be discussed in detail. Included will be comments in regard to therapy and means of protecting the skin from the harmful effects of sunlight.

SUNBURN AND SUN TAN

The sunburn-producing waves of sunlight are from 2,900 to 3,170 Å in length, with the greatest effect occurring at 2,970 Å. Erythema may be produced artifically by either a hot quartz or a cold quartz ultraviolet lamp.[3] Erythema that is produced by a cold quartz ultraviolet lamp is caused by wave lengths of 2,537 Å. Such wave lengths are not present at the surface of the earth since atmosphere, particularly ozone, filters out all wave lengths shorter than 2,900 Å.[4] Although the mechanism through which ultraviolet produces erythema is not definitely established, it is theorized that injured cells liberate leukotaxine or other substances which induce inflammation.

Following an exposure to sunshine, the degree of cutaneous reaction is dependent upon several factors, the most important of which are the duration and intensity of exposure and the degree of natural protection afforded by the epidermis. It is interesting that there are more erythemogenic wave lengths present in sunlight at midday than early or late in the day, in southern than northern latitudes, and at high altitudes because of less atmospheric absorption. Reflection, such as occurs from snow and at the seashore, will also increase the intensity of an exposure. On the other hand, smoke, clouds, shade, sun tan, and clothing provide protection.

The stratum corneum absorbs and scatters ultraviolet and this outermost layer of the skin will stop a major portion of the incoming ultraviolet energy. Natural protection from sunlight is determined primarily by two factors, one being the thickness of the stratum corneum and the other the amount of melanin pigment in the skin. Melanin is found in the stratum corneum as well as in the basal cell layer of the epidermis, and Thompson[5] has shown that the number of melanin granules in the stratum corneum significantly influences its ability to absorb ultraviolet light. An albino will sunburn after a minimal exposure, whereas colored skin is remarkably resistant to sunburn. It is apparent that susceptibility to sunburn is, to a great extent, inversely proportional to the amount of pigment in the epidermis.

A sun tan can significantly increase an individual's resistance to sunburn. There are three separate and different processes involved in the production of a sun tan. Within a few minutes after irradiation there is oxidation of many melanin granules, and oxidized melanin is darker than reduced melanin. Second, melanin granules are dispersed following irradiation, and skin in which the melanin granules are dispersed is darker than it was previously. This dispersion reaction is a delayed effect, and its maximum is not reached for several days following irradiation. These first two reactions are caused by relatively long ultraviolet wave lengths. Third, ultraviolet wave lengths between 2,800 and 3,100 Å have the ability to cause melanocytes to produce new melanin. New melanin is not observed until after 48 hours, and melanin production does not reach its maximum for approximately 19 days.[6] Freckling is a form of tanning which occurs most frequently in red or reddish-brown-haired individuals. In a freckle there is an unusually large amount of pigment without an actual increase (possibly a decrease) in the number of melanocytes.[7]

ACTINIC DEGENERATION AND CARCINOGENESIS

Exposed skin is subjected to a large amount of irradiation from sunlight during a lifetime. Factors which influence the amount of irradiation a person receives include occupation, geographic location, clothing, recreational activities, etc. Most skin cancers occur in light-exposed areas, and sunlight is considered to be the major cause of skin cancer in man. The high incidence of skin cancer in sailors was noticed by Unna,[8] and his conclusion that sunlight was a major cause of skin cancer has been adequately confirmed.

Premalignant and malignant skin tumors develop on sunlight-damaged skin. Fair-skinned individuals who have been exposed to large amounts of sunshine manifest the highest incidence of skin cancer. Macdonald[9] has compiled interesting statistics showing that skin cancer is more common in areas of intense sunlight and in light-skinned races. Wave lengths between 2,900 and 3,341 Å are the carcinogenic wave lengths of the ultraviolet spectrum. In experimental animals the precancerous period varies inversely with the daily dose of irradiation, and carcinogenesis, once initiated, will proceed without further exposure to ultraviolet.[2] A similar situation seems to exist in man, although animal studies cannot always be transferred to humans.

The amount of radiant energy necessary to precipitate carcinogenesis in albino mice is relatively small, with 63 to 84 $\times$ 10^7 ergs per square centimeter required.[10] In a series of experiments with Swiss strain albino mice we have been able to demonstrate that topical sun screens greatly reduce experimental ultraviolet carcinogenesis.[11] Since 8-methoxypsoralen can be used to accelerate carcinogenesis, some groups of animals were given the drug for this purpose. In addition to the finding that topical sun screens can inhibit ultraviolet carcinogenesis, it was discovered that orally administered antimalarials and 3-benzoyl-4-hydroxy-6-methoxy-benzenesulfonic acid could decrease the incidence of skin tumors.[12] Control animals developed approximately twice as many cancers as those receiving these drugs in their diet.

Sunlight-induced carcinogenesis has been adequately emphasized; however, there has been relatively little comment in the literature in respect to the role that sunlight plays in aging of the skin. After years of exposure to sunlight, the skin manifests degenerative changes, both clinically and histologically. Among the visible changes related to chronic exposure to sunlight are a dry, coarse, and leathery appearance of the skin, pigmentary changes, and wrinkling. Because pigment provides excellent protection from sunlight-induced degeneration, elderly Negroes often manifest a remarkably youthful appearance. A person who spends most of his life outdoors, such as a farmer, will often appear considerably older than a white-collar worker of similar age. Such differences would be even more noticeable if it were not for graying, skeletal changes, and other phenomena that accompany aging. If one examines an elderly person's skin, he will note a remarkable difference between the light-exposed areas and the regions protected by clothing. A 70-year-old woman with an aged, wrinkled, and withered face will often have skin similar to that of a young woman on her chest.

PHOTOSENSITIVE DERMATOSES

Photosensitivity is not an unusual phenomenon. Although the majority of patients who manifest photosensitivity have no systemic disease, photosensitivity may be a cutaneous manifestation of systemic disease, with pellagra, porphyria, and lupus erythematosus as interesting examples. The most common photosensitive dermatosis is the polymorphic light eruption, which is idiopathic in origin. It is important to know that drugs, such as the sulfonamides and phenothiazine derivatives, can induce photosensitivity. A photosensitive individual will manifest a dermatitis which is limited to, or more severe in, light-exposed areas, and most patients are conscious of the fact that sunlight aggravates their condition. Since serious systemic diseases and a number of drugs can produce photosensitivity, it is quite important that every patient who presents a photosensitivity dermatosis be carefully evaluated. A detailed history, thorough physical examination, and appropriate laboratory studies will usually reveal information that will enable the physician to establish a correct diagnosis.

A. Systemic Diseases With Photosensitivity.—

Porphyria: There are two main types of porphyria, erythropoietic and hepatic. The hepatic type is divided into porphyria cutanea tarda and acute intermittent porphyria. Photosensitivity of a varying degree occurs in porphyria cutanea tarda and in the erythropoietic variety. Porphyrins absorb ultraviolet between 3,000 and 6,500 Å, with major absorption bands between 3,000 and 4,100 Å, and between 5,060 and 6,020 Å, depending upon the type of porphyria.[13]

The erythropoietic type is congenital, and photosensitivity is a prominent feature of the syndrome. Males are affected approximately twice as often as females. The abnormality is first discovered in infancy or in early childhood, with the first symptoms being pink or reddish urine and photosensitivity. Erythematous papules and vesicles develop on the face, ears, and uncovered portions of the extremities following exposure to sunlight. Brown stains appear on the teeth, and at autopsy bones will be found to be impregnated with porphyrins. Melanosis and hypertrichosis may occur. Late complications include splenomegaly, anemia, and osteoporosis. The clinical diagnosis can be confirmed by finding uroporphyrin, primarily uroporphyrin I, in the urine. There is no satisfactory treatment, and these children must be protected from sunlight.

The hepatic type of porphyria may manifest itself as the acute intermittent type, porphyria cutanea tarda, or a mixture of these two varieties. In acute intermittent porphyria a number of abdominal or nervous system symptoms may occur. Abdominal pain can be severe enough to suggest an acute abdominal crisis, and it is not unusual for such patients to have undergone abdominal surgery on one or more occasions. The urine may be port wine in color during acute attacks; however, it is often clear. Diagnosis depends upon the findings of porphobilinogen in the urine. Photosensitivity does not occur in the pure, acute intermittent type of porphyria. Porphyria cutanea tarda is characterized by bullae on the hands, melanosis, and a bleary-eyed, dissipated appearance. Some of these patients have eczematous lesions on the face and ears, and vesicles and bullae can occur in these areas. Trauma often precipitates the development of a

bullous lesion. Liver function tests may reveal liver damage and, in many instances, patients with porphyria cutanea tarda are alcoholics. Uroporphyrin is found in the urine, although repeated examinations may be necessary before the porphyrins can be demonstrated. Wood's light will produce a pink fluorescence in urine when porphyrins are present. Mixed porphyria may present features of either type or may manifest the picture of acute porphyria on one occasion and cutanea tarda at another time.

Pellagra: Pellagra is a vitamin deficiency disease in which the prime deficiency is of nicotinic acid. Although it is uncommon today, occasional cases of pellagra are found. In our area, pellagra is most frequently found in individuals with psychiatric disorders, recluses, alcoholics, or elderly widowers who have an improper food intake. The well-known triad of diarrhea, dementia, and dermatitis is occasionally seen; however, most cases are less typical and are borderline in nature. Photosensitivity, with the primary involvement on the backs of the hands, is somewhat characteristic. The face and other exposed areas may manifest photosensitivity phenomena. The tongue is often red, and gastrointestinal disturbances are common. Nervous system symptoms may also be present, and include fatigue, insomnia, anorexia, vertigo, neuritic symptoms, headache, apprehension, etc. In advanced cases, psychoses and more serious neurological disturbances occur.

Lupus erythematosus: This disease may be divided into discoid lupus erythematosus and disseminated lupus erythematosus. In the discoid variety, skin lesions are the only manifestation of the disease, while, in the disseminated type, various organ systems may be involved. The systemic disease may be so mild that the disease is subacute in character, or the symptoms may be of such magnitude as to produce an acute illness that frequently terminates in death. Although it is rare for discoid lupus erythematosus to change into disseminated lupus erythematosus, the diseases are closely related, and probably represent variations in the same disease process.[14-17]

In discoid lupus erythematosus the skin lesions are sharply demarcated, have an erythematous border, and manifest central atrophy and scarring. Follicular plugging is often evident. Skin lesions of the classic discoid variety may be present in disseminated lupus erythematosus; however, lesions of acute lupus erythematosus are usually more inflammatory, and atrophy with scarring is not as prominent a feature. Histopathologic changes are quite similar in both varieties.

Many patients with either discoid or disseminated lupus erythematosus are impressed with the adverse effect that sunlight has on the course of the disease. Lesions are located primarily on light-exposed areas of the body, the cheeks being the predominant location, and exposures to sunlight may produce serious exacerbations of disseminated lupus erythematosus.[1] Fatigue, fever, and arthralgia are among the most constant systemic symptoms. Leukopenia, positive L.E. cell test, increased sedimentation rate, positive serologic test, hematuria, and proteinuria are among the more significant laboratory findings.

B. Congenital.—

Xeroderma pigmentosum: This disease is a rare, familial condition that appears during childhood. It is characterized by a type of photosensitivity which

results in actinic degeneration of the skin, with the subsequent early development of premalignant and malignant skin tumors. Light-exposed areas develop atrophy, telangiectasia, and hyperpigmentation, upon which numerous premalignant and malignant tumors develop. Photophobia, lacrimation, keratitis, and opacities of the cornea occur. Because of the numerous squamous cell carcinomas that continue to develop, these patients eventually develop metastases, and very few live past adolescence.

C. Photodynamic Reactions.—A photosensitizing agent is necessary for the production of a photodynamic dermatosis. The photosensitizing chemical absorbs radiant energy and acts somewhat like a catalyst. Blum[1] believes that the absorbed photons of light give the absorbing molecules a higher energy content, and subsequently this excess energy is transferred from the photosensitizer to the substrate. The transferred energy produces oxidative reactions in the substrate which damage cellular structures. Photosensitizers usually absorb only specific wave lengths, and after the release of energy they can accept more energy. Some photodynamic dermatoses are photoallergic in origin, and others are phototoxic in origin. A phototoxic reaction is not due to any altered reactivity of the individual, but to a pharmacodynamic reaction which depends upon the absorption characteristics of the particular agent involved. Sulfonamides, phenothiazines, meadow grass, buckwheat, coal tar, figs, limes, parsley, and certain perfumes are examples of agents that can produce photodynamic responses.

D. Other Photosensitivity Reactions.—Photoallergic diseases are a form of physical allergy. It has been theorized that specific metabolites are liberated after the skin is exposed to certain physical agents, and rare individuals develop a true allergic sensitivity to these metabolites.[18] True allergy has been shown in solar urticaria by the passive transfer technique; however, polymorphic light-sensitive eruptions have not been proved to be the result of an allergic reaction. Lamb and associates[19] have classified polymorphic light-sensitive eruptions into four clinical variants: (1) plaquelike type, (2) contact, eczematous type, (3) papular and prurigo-like type, and (4) erythematous type. The plaquelike type has erythematous, slightly infiltrated, plaquelike lesions in light-exposed areas, whereas, the contact, eczematous type is eczematous in nature. The third reaction pattern consists of papules, papulovesicles, and small erythematous wheals. Erythema perstans–like lesions characterize the fourth morphologic variant. Polymorphic light-sensitive eruptions usually develop in the spring or summer and disappear in the fall or winter. Patients will often state that sunlight causes, or at least aggravates, the condition. The etiology of these eruptions remains obscure. It is interesting that the so-called "daylight" fluorescent lights can produce polymorphic light-sensitive eruptions.[20]

Hydroa aestivale is a photosensitivity dermatosis that occurs almost exclusively in children and adolescents. Small vesicles and excoriated papules are found on the cheeks, nose, and ears. The eruption appears in the summer and clears during the winter. Hydroa aestivale usually disappears around puberty.

THERAPY

Excellent protection from sunburn can be provided by sun screens. Several families of compounds have been used as sun screens, with para-aminobenzoates, anthranilates, cinnamates, pyrrones, benzimidazoles, carbazoles, naphthosulfonates, and quinine disulfate being most frequently employed. Rothman[21,22] found that a 15 per cent para-aminobenzoic acid cream would increase the minimal erythema threshold from 50 to 100 fold. Knox and co-workers[23] showed that the benzophenones are excellent sun screens and equal to, or superior to, other sun-screening agents. This investigation also revealed that the best currently available commercial sun screens will increase the minimal erythema threshold only 4 to 6 times. In comparative evaluations, 10 per cent concentrations of 3-benzoyl-4-hydroxy-6-methoxy-benzene-sulfonic acid and para-aminobenzoic acid in vanishing cream were found to increase the minimal erythema dose by approximately 100 fold. Benzophenones have an exceptionally wide absorption spectrum that extends from very short wave lengths almost to visible light; therefore, they absorb the wave lengths that produce erythema, pigmentation, and many photosensitivity reactions. The better commercial sun-tanning preparations should enable a sun bather to receive 4 to 6 times as much exposure without developing a sunburn, whereas special formulations can provide near-perfect protection.

Vol. 13
1961

Para-aminobenzoic acid and its derivatives do not absorb the longer sun-tanning wave lengths, and are the active agents in most sun tan preparations. For individuals who need protection from sunlight and are not anxious to obtain a tan, a benzophenone should be the sun screen of choice. Stegmaier[24] found that a benzophenone was more effective in inhibiting 8-methoxypsoralen-accelerated pigmentation than was para-aminobenzoic acid. From the evidence available, it appears that a sun screen of the benzophenone class would be of value for individuals who freckle, patients with chloasma, and people who prefer their natural color to sunlight-induced hyperpigmentation.

Physical sun screens (sunshades) produce their effect by scattering light rather than absorbing specific wave lengths. Titanium dioxide, talc, kaolin, zinc oxide, and bentonite can be used as physical sun screens, and occasionally these agents are included in formulations along with chemical sun screens. Window glass filters out the erythema-producing wave lengths, and therefore protects from sunburn. A plain lubricating oil or cream will provide some protection, and petrolatum, particularly red petrolatum, delivers considerable protection.

The most important factor in the prevention and management of harmful effects of sunlight on the skin is protection of the individual from sunshine. Protective measures include a detailed explanation to the patient about how sunlight is causing his disease, advice to remain out of the sun as much as possible, and recommendations in regard to the wearing of long-sleeved garments, a hat, and a sun screen. A prescription for 10 per cent para-aminobenzoic acid in a vanishing cream base can be filled in most pharmacies. Wide-spectrum sun screens of the benzophenone type are not yet available through drugstores; however, many patients are benefited by the use of proprietary preparations, such as Skolex*

*Skolex, The J. B. Williams Co., Division of Pharmaceuticals Inc., Newark, N. J.

and A-Fil.* Most patients do not realize that even a limited amount of exposure may be harmful; therefore it must be emphasized that protective measures should be employed for even minor outings, such as a short shopping trip. Although the above-mentioned sun screens do not absorb all of the photosensitizing wave lengths, many of the previously described sunlight-induced conditions will be benefited by their use.

Following Page's[25] report in 1951, that quinacrine was effective in the treatment of discoid lupus erythematosus, the antimalarials have been widely employed in the treatment of diseases in which sunlight is a causative factor. The antimalarials have their greatest usefulness in discoid lupus erythematosus and in polymorphic light-sensitive eruptions. These drugs are also of value in the treatment of subacute lupus erythematosus and in acute lupus erythematosus after the disease has been controlled by systemic steroids. At this point it should be emphasized that antimalarials are contraindicated in porphyria for they may produce serious, untoward effects in this disease.[26] These agents can also be administered for other conditions such as actinic keratoses.[12]

Chloroquine has been the most widely used antimalarial, and it is usually effective in small doses, such as one 250 mg. tablet a day. Many dermatologists prefer to start their patients on 2 tablets per day for one to two weeks, and then reduce the dose to one 250 mg. tablet per day. The discovery that the antimalarial drugs are effective for these diseases was a major contribution to medical knowledge.

The mechanism of action of the antimalarial drugs in lupus erythematosus and light-sensitive diseases has not been established. The chemical tends to localize in nucleated cells with the greatest concentration in the liver and spleen. Lesser quantities are found in the kidneys, lungs, and skin.[27,28] At first they were thought to act by screening certain ultraviolet wave lengths, thereby diminishing the adverse effects produced by sunlight. Cahn and his collaborators[29] have shown that abnormal, but not normal, skin responses to ultraviolet light are altered. Also, Hall[30] successfully treated 3 cases of solar dermatitis with chloroquine without interfering with the tanning effect of sunshine. According to Shaffer and his associates[31] the quantity of antimalarials present in the epidermis is not enough to act as a significant physical barrier to the passage of ultraviolet light, since spectrographic studies performed by them disclosed that the drugs had no perceptible effect on ultraviolet absorption by epidermis.

It has been suggested by Kurnick[32] that these compounds may have the ability to combine with nucleic acids to prevent the deposition of depolymerized nuclear protein in tissues. Lerner and Lerner[33] believe that the action of quinacrine and chloroquine may be the result of an effect on flavin enzymes.

Although most sunlight-induced diseases can be successfully managed with education, protection, and antimalarials, some cases are benefited by additional measures or a different approach. Each case should be individualized and treated accordingly. Topical and systemic steroids, testosterone, antihistamines, gold, and other agents may be useful in selected cases.

*A-Fil, Texas Pharmacal Co., San Antonio, Texas.

SUMMARY

Most of the significant diseases in which sunlight is a causative or contributing factor were discussed in detail. Conditions receiving greatest emphasis were sunburn, photosensitivity, aging, and carcinogenesis. Patient education, protection with clothing and sun screens, and the antimalarial drugs are considered to be the basis of therapy.

REFERENCES

1. Blum, H. F.: Photodynamic Action and Diseases Caused by Light, New York, 1941, Reinhold Publishing Corporation.
2. Blum, H. F.: Carcinogenesis by Ultraviolet Light: An Essay in Quantitative Biology, Princeton, N. J., 1959, Princeton University Press.
3. Coblentz, W. W.: Sources of Ultraviolet and Infrared Radiation Used in Therapy, J.A.M.A. **132**:378, 1946.
4. Sanderson, J. H., and Hallburt, E. O.: *In* Hollaender, E. H., editor: Radiation Biology, New York, 1955, McGraw-Hill Book Co., Inc.
5. Thompson, M. L.: Relative Efficiency of Pigment and Horny Layer Thickness in Protecting the Skin of Europeans and Africans Against Solar Ultraviolet Radiation, J. Physiol. **127**:236, 1955.
6. Rothman, S.: Physiology and Biochemistry of the Skin, Chicago, 1954, University of Chicago Press.
7. Breatnach, A. S.: Melanocyte Distribution in Forearm Epidermis of Freckled Human Subjects, J. Invest. Dermat. **29**:253, 1957.
8. Unna, P. G.: Orth's Lehrbuch der speziellen pathologishen anatomie, Berlin, 1887, Hirschwald, p. 719.
9. Macdonald, E. J.: The Epidemiology of Skin Cancer, J. Invest. Dermat. **32**:379, 1959.
10. Griffin, A. C., Hakim, R. E., and Knox, J. M.: The Wave Length Effect Upon Erythemal and Carcinogenic Response in Psoralen Treated Mice, J. Invest. Dermat. **31**:289, 1958.
11. Knox, J. M., Griffin, A. C., and Hakim, R. E.: Protection From Ultraviolet Carcinogenesis, J. Invest. Dermat. **34**:51, 1960.
12. Knox, J. M., Griffin, A. C., and Hakim, R. E.: The Effect of Chloroquine on the Erythematous and Carcinogenic Response to Ultraviolet Light, A.M.A. Arch. Dermat. **81**:570, 1960.
13. Kesten, B. M., and Slatkin, M.: Diseases Related to Light Sensitivity, A.M.A. Arch. Dermat. **67**:284, 1953.
14. Huff, S. E., Taylor, H. L., and Keys, A.: Observations on the Peripheral Blood Flow in Chronic Lupus Erythematosus, J. Invest. Dermat. **14**:21, 1950.
15. Goldman, L.: New Developments in the Treatment of Discoid Lupus Erythematosus, M. Clin. North America, p. 391, March, 1951.
16. Montgomery, H.: Pathology of Lupus Erythematosus, J. Invest. Dermat. **2**:343, 1939.
17. Michelson, H. E.: Acute Lupus Erythematosus, M. Clin. North America, p. 399, March, 1951.
18. Baer, R.: Allergic Dermatoses Due to Physical Agents, New York University Press, Philadelphia, 1956, by J. B. Lippincott Company, Distributors.
19. Lamb, J. H., Shelmire, B., Cooper, Z., Morgan, R. N., and Keaty, C.: Solar Dermatitis, A.M.A. Arch. Dermat. & Syph. **62**:1, 1950.
20. Lamb, J. H., Jones, P. E., Rebell, G., and Alston, H. D.: Sensitivity to Fluorescent (Blue-Green) Light, A.M.A. Arch. Dermat. **77**:519, 1958.
21. Rothman, S., and Rubin, J.: Sunburn and Para-aminobenzoic Acid, J. Invest. Dermat. **5**:445, 1942.
22. Rothman, S., and Henningsen, B. A.: The Sunburn Protecting Effect of Para-aminobenzoic Acid, J. Invest. Dermat. **9**:307, 1947.
23. Knox, J. M., Guin, J. D., and Cockerell, E. G.: Benzophenones. Ultraviolet Light Absorbing Agents, J. Invest. Dermat. **29**:435, 1957.
24. Stegmaier, O. C., Jr.: Discussion of Knox, Griffin, and Hakim.[11]
25. Page, F.: Treatment of Lupus Erythematosus With Mepacrine, Lancet **2**:755, 1951.
26. Davis, M. J., and Ploeg, D. E. V.: Acute Porphyria and Coproporphyrinuria Following Chloroquine Therapy, A.M.A. Arch. Dermat. **75**:796, 1957.
27. Dearborn, E., Kelsey, F., Oldham, F., and Geiling, E.: Studies on Antimalarials. The Accumulation and Excretion of Atabrine, J. Pharmacol. & Exper. Therap. **78**:120, 1943.
28. Barlow, O., Auerbach, M., and Rivenburg, H.: Studies of the Pharmacology of Atabrine on Mice, Rats, and Dogs, J. Lab. & Clin. Med. **39**:20, 1945.

29. Cahn, M., Levy, E., and Shaffer, B.: The Effect of Chloroquine Phosphate in Modifying Reactions to Ultraviolet Light, J. Invest. Dermat. **26**:201, 1956.
30. Hall, F. A.: Discussion of Lamb, J., Jones, P., and Maxwell, T.: Solar Dermatitis, A.M.A. Arch. Dermat. **75**:171, 1957.
31. Shaffer, B., Cahn, M., and Levy, E.: Absorption of Antimalarial Drugs in Human Skin, J. Invest. Dermat. **30**:341, 1958.
32. Kurnick, N.: A Rational Therapy of Systemic Lupus Erythematosus, A.M.A. Arch. Int. Med. **97**:562, 1956.
33. Lerner, M. R., and Lerner, A. B.: Dermatologic Medications, Chicago, 1954, The Year Book Publishers, Inc.

Vol. 13
1961

DIAGNOSIS AND TREATMENT OF LUPUS ERYTHEMATOSUS, DERMATOMYOSITIS, AND SCLERODERMA, WITH EMPHASIS ON CUTANEOUS FINDINGS

R. K. WINKELMANN, M.D., PH.D.
ROCHESTER, MINN.

*From the Section of Dermatology, Mayo Clinic and Mayo Foundation**

(Received for publication Nov. 22, 1960)

THE term "collagen disease" now is considered slightly disreputable when used inclusively to designate systemic lupus erythematosus (S.L.E.), dermatomyositis, and scleroderma. No basic pathogenetic link has been found between these diseases; the "fibrinoid" change, once ascribed to alterations in the connective-tissue ground substance, now is attributed to actual fibrin. A circumstance that unites these three conditions is the presence of characteristic cutaneous lesions; in many instances, these are identical clinically. While these changes are not universally present, the manifold expressions of these diseases in the skin lead the trained clinician toward the correct diagnosis. These diseases were described originally through their dermatological changes, and the visual patterns remain a keystone in their diagnosis.

LUPUS ERYTHEMATOSUS

Chronic discoid lupus erythematosus (D.L.E.) is a papulosquamous eruption that terminates in atrophy and telangiectasia. Keratotic plugging of the follicles is a clinical and histologic feature. The typical location is on the malar prominence and over the nose, but lesions of the eyelids, lips, ears, and scalp are expected. An extension of the process onto the glabrous skin, particularly on the exposed areas of the extremities and trunk, has given rise to the term "generalized discoid lupus erythematosus." Aggravation by solar exposure is usual. Lesions of the mucous membranes occur in 20 per cent of cases in which the disease is extensive and consist of bluish-white patches with peripheral telangiectasia and occasional ulceration. The eruption may occur at any age. An example of D.L.E., commencing at 2 years of age and still present at the age of 15 years, is seen in Fig. 1,*a*. The response to 4 weeks of treatment with amodiaquin (Camoquin) hydrochloride is illustrated in Fig. 1,*b*. While such childhood cases are unusual, they typify the chronic course of the disease. Scarring is usual, and it may produce disfigurement and, in the scalp, permanent alopecia. Epitheliomas rarely occur in such chronic scars.

*The Mayo Foundation, Rochester, Minn., is a part of the Graduate School of the University of Minnesota.

Recognition of the L.E. phenomenon in the blood has widened the understanding of cases that in prior years would have been diagnostic enigmas. It is now possible to classify as S.L.E. instances of obscure disease of the renal, articular, hematopoietic, and nervous systems. Similarly, most instances of disease formerly considered as subacute L.E. are now recognized as S.L.E. Even cases of generalized D.L.E. are now recognized as true S.L.E. at the start of observation, rather than being considered as D.L.E. that has been converted to S.L.E. This increase in the number of cases of D.L.E. now being classified as S.L.E. does not indicate that the conversion rate is higher but rather that the ability to define the clinical state has improved. The older estimates of a 20 to 30 per cent conversion rate of D.L.E. to S.L.E. should be ignored. A conversion rate of 6 per cent, based on 132 cases studied by Montgomery and McCreight,[1] is confirmed by present clinical experience.

Fig. 1.—*a*, Chronic discoid lupus erythematosus present for 13 years in a 15-year-old girl. *b*, Response to use of Camoquin after 4 weeks.

The specific lesions of S.L.E. are similar clinically and histopathologically to those of D.L.E. Cutaneous lesions occurred in all but 16 of 105 cases reviewed by Harvey and associates.[2] The acute onset of S.L.E. tends to emphasize the edematous and erythematous components of the process rather than the atrophic and squamous factors but, with time, as Montgomery[3] has shown, the acute cutaneous lesions of S.L.E. develop into typical chronic discoid lesions. The nonspecific lesions of S.L.E. may begin with chronic persistent erythema, which may be present on the palms, over the extensor portions of the body on the light-exposed areas, and on the dorsum of the fingers between the joints. This persistent erythema may proceed to a reticulated, atrophic, telangiectatic, and hyperpigmented state termed "poikiloderma." Petechiae with erythema may occur on the volar surfaces of the finger tips (Fig. 2,*a*) and about the nail folds as a

typical indication of S.L.E. True purpura may occur from a thrombocytopenic state accompanying S.L.E. or as a result of angiitis of the skin. Livedo reticularis and infarctive ulcers are typical of cutaneous angiitis in L.E. (Fig. 2,*b*). These lesions are usually found in the dependent portions of the body, but they may occur in any location. While the superficial ulceration of angiitis heals readily, chronic indolent ulcers of the deep tissues also occur and heal slowly. Raynaud's phenomena are present in 10 per cent of the cases recognized today and should be anticipated.

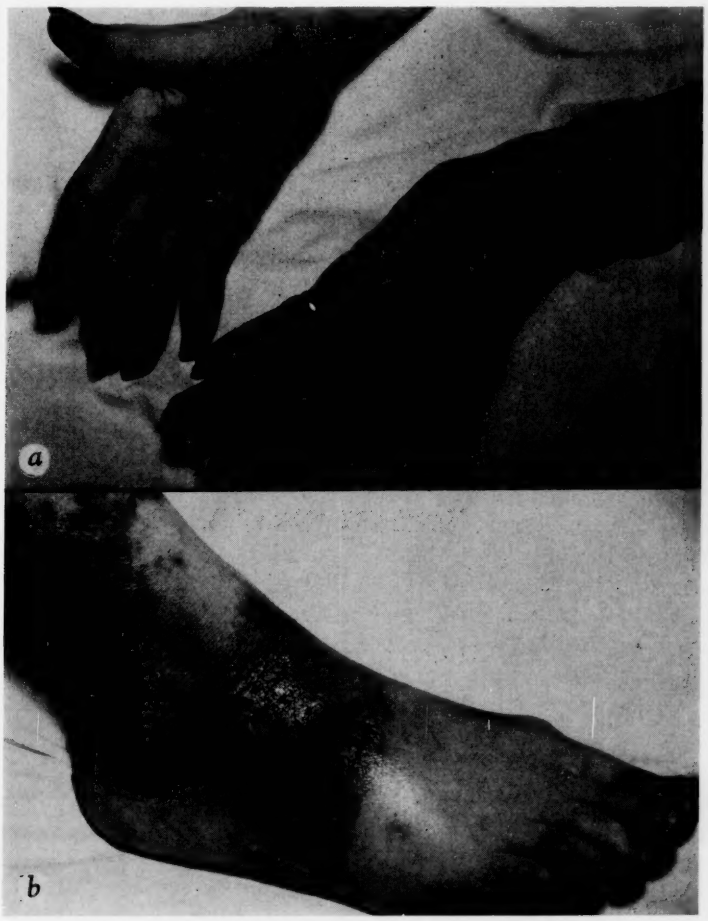

Fig. 2.—*a*, Petechial erythema about the finger tips in systemic lupus erythematosus. *b*, Another case, showing the reticulated pattern of petechial livedo and infarctive ulceration in the angiitis of systemic lupus erythematosus. Note the superficial quality of the ulcer and its jagged edges.

The acceptance of the L.E. phenomenon as indicative of the L.E. state has emphasized a range of disease processes that may be mentioned here. One cutaneous entity recently recognized in this category is Sjögren's syndrome, in which dryness of mucous membranes is associated with arthritis. Another type

of arthritis associated with positive results of the L.E. clot test occurs in Felty's syndrome, in which the only major cutaneous sign is chronic ulceration of the lower extremities. This test gives positive results in many acutely ill patients with rheumatoid arthritis; in general, these patients have true S.L.E. rather than a rheumatoid state converted to S.L.E. by the use of steroids. Positive results also are seen in many acutely ill patients with chronic ulcerative colitis and arthritis, and they occur occasionally in periarteritis nodosa and dermatomyositis, as well as in reactions induced by use of hydralazine (Apresoline) and hydantoin. An important addition to knowledge regarding this phenomenon is the recognition of positive results of the L.E. clot test in posthepatitic cirrhosis.[4] Necropsy in these cases may not reveal any evidence of S.L.E. with the cirrhosis despite the fact that, during life, signs of hepatic disease were associated with arthritis, positive results of the L.E. clot test, and hyperglobulinemia.

All these states now considered by various authors as representing a L.E. reaction have certain findings in common. In most instances, these changes include hyperglobulinemia, an increased erythrocytic sedimentation rate, and autoimmune phenomena that depress the formed elements of the blood and bone marrow, producting leukopenia and occasionally thrombocytopenia and hemolytic anemia. It is too early to say that only those conditions associated with pathologic proof of L.E. can be accepted as L.E. At present, it appears that only time and observation will allow one to make these fine distinctions.

The treatment of D.L.E. consists principally in administration of the 4-amino-quinoline type of antimalarial compounds. Quinacrine (Atabrine), chloroquine, hydroxychloroquine (Plaquenil), and Camoquin have been given successfully. Combinations of these drugs are on the market and are also useful; one such combination is Triquin. I usually select chloroquine first, followed by Camoquin if the condition proves resistant. Because of the toxicity of these drugs, they should not be used in large doses for long periods. In general, their in vitro and in vivo effect on the L.E. phenomenon can be accomplished with relatively small doses, even in sick patients; 250 mg. of chloroquine and 200 mg. of Camoquin given daily should be the maximal maintenance dose.

The treatment of S.L.E. involves the use of steroids when acute fever, joint pain, pleurisy, or abdominal pain, or other acute symptoms are present. Hydrocortisone or methylprednisolone (Medrol), in moderate doses, is usually effective in suppressing the inflammatory response. The coincident administration of an antimalarial drug will depress the L.E. phenomenon in time, and patients have been treated successfully first with steroids and antimalarial agents and then with antimalarial drugs alone for 5 years after the diagnosis of S.L.E. has been made. A number of my patients have experienced successful pregnancies on such a regimen, which is a further test of its usefulness. Contrary to the opinion of Muehrcke and associates,[5] I believe that it is possible to cause reversion of the renal findings in L.E., and I do not think that clinical or pathologic criteria exist at present for prior selection of those patients who might respond. The presence of renal damage always has been considered as a serious prognostic sign, but it should be emphasized, however, that vigorous treatment is worth while, even in these cases, for an occasional patient will respond satisfactorily.

DERMATOMYOSITIS

Dermatomyositis is a primary inflammatory disease that may attack principally skin or muscle or both. When muscles are involved alone, it is usually known as "polymyositis." The onset may be slow or rapid. It may follow an infection of the upper part of the respiratory tract. The most notable clinical observations usually are related to muscular weakness and the development of certain cutaneous lesions. The cutaneous changes are not specific but are rather indications for the diagnosis of dermatomyositis.

In the acute phase of dermatomyositis, massive edema of the face, eyelids, and involved skin may be present (Fig. 3). Associated with this may be mild to intense erythema, and a livid or violaceous (heliotrope) color may be present over the eyelids. Chronic persistent symmetric erythema may be seen in this eruption. This may follow the acute edematous and erythematous phase, or it may commence on the hands, forearms, face, or upper part of the thorax. It is notable for its location over the extensor surfaces of joints and pressure points.

ol. 13
1961

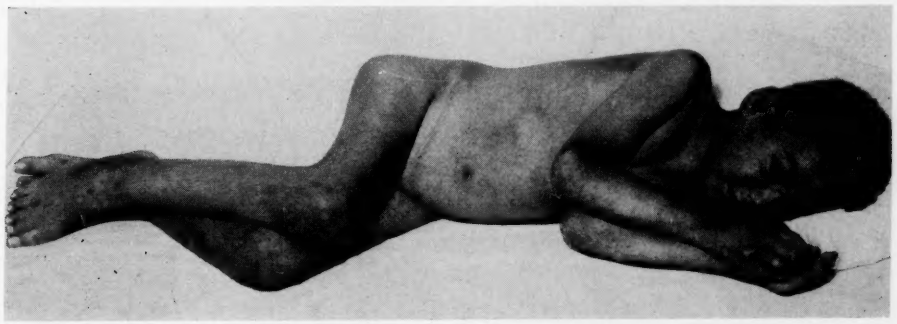

Fig. 3.—Edema and erythema of the face, eyelids, and extensor surfaces of the joints in a young boy with acute dermatomyositis. Muscular atrophy is apparent.

If this eruption persists for a long period, it becomes poikilodermatous (Fig. 4). The original cases of dermatomyositis were described as poikilodermatomyositis. The occurrence in this disease of telangiectasia and atrophy, together with hyperpigmentation, has presented a clinical problem with regard to distinguishing this condition from S.L.E. Raynaud's phenomena occur in approximately 20 per cent of the cases. With progression of the disease, scarring and contraction may occur.

Changes resembling those of acrosclerosis involving the hands, forearms, face, and upper part of the thorax were noted in one fourth of a group of cases studied at the Mayo Clinic by my associates and me.[6] As a result of this study of calcinosis cutis in dermatomyositis, I think that most calcinosis universalis found in the central portions of the trunk results from prior episodes of dermatomyositis. Calcinosis cutis has developed in 40 per cent of the children with dermatomyositis studied at the Clinic. Almost all patients with calcinosis cutis are female. The extrusion of these calcific masses may cause great discomfort and high fever, and may pose a serious problem in the management and rehabilitation of the patient.

The diagnosis of dermatomyositis rests equally on electromyography, measurement of serum transaminase, and biopsy of muscle. Clinical evaluation of muscular function as well as assessment of the general body status is important. The differences in laboratory findings between dermatomyositis and S.L.E. are remarkable enough to give ready differentiation in most cases.[7] As already indicated, an occasional case of dermatomyositis resembles acrosclerosis on one hand or S.L.E. on the other.

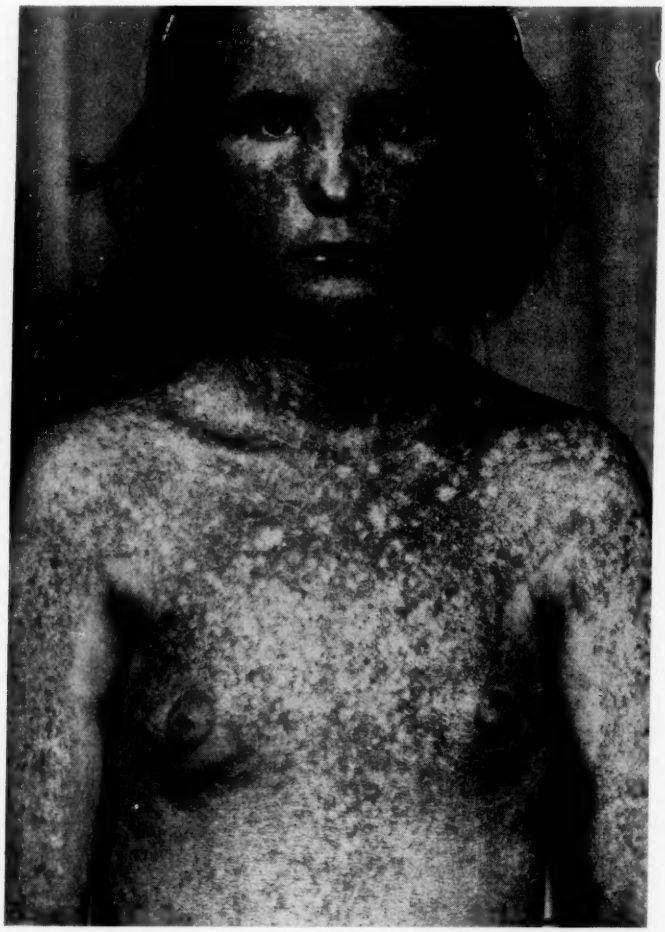

Fig. 4.—Extensive poikilodermatous changes in a young girl with dermatomyositis.

The relationship of carcinoma to dermatomyositis has been emphasized in recent years.[8] Some authorities state that laparotomy is advisable in all cases of dermatomyositis. Contrary to this opinion is the conclusion of a retrospective study of 270 cases of dermatomyositis in which only a 6 per cent incidence of

carcinoma was discovered.[9] No cases of concurrent malignant tumor and dermatomyositis have been reported in children, so the presence of a malignant lesion is not a sine qua non for the diagnosis of dermatomyositis. It is necessary to examine every adult patient with dermatomyositis for signs of tumor, but one must be satisfied to find it in a reasonably small number of patients.

The treatment of dermatomyositis at present is not specific but it is possible to reduce symptoms and restore function to the inflamed muscle.[10] Steroids should be given to suppress the signs and symptoms of inflammation. When the values for serum transaminase and the clinical response indicate regression of inflammation, physical therapy may be started.[11] Usually, it is only after administration of small maintenance doses of steroids for 6 to 9 months that changes in the electromyogram begin to occur. Chelation with edathamil disodium (EDTA) has proved useful in some patients, and the use of anabolic steroid agents such as norethandrolone (Nilevar) or testosterone apparently has been helpful in some instances. The disease may continue for 2 to 4 years, and maintenance doses of steroids may be required for such periods. It is important to establish suppression of muscular inflammation with doses of steroids small enough to avoid the well-known side effects of adrenocortical hormones.

SCLERODERMA

Scleroderma, or hardening of the skin, encompasses several disease processes. A useful classification recognizes (1) localized forms, including morphea, linear scleroderma, and the hemiatrophic stage, and (2) generalized systemic sclerosis, which includes acrosclerosis and diffuse scleroderma. It is difficult at times to differentiate these conditions from such cutaneous diseases as scleredema and lichen sclerosus et atrophicus, and from the surface sclerosis that may be found in progressive S.L.E., rheumatoid states, and dermatomyositis. Diagnostic efforts to exclude other systemic diseases always must be made.

Localized scleroderma has been reviewed by Christianson and associates.[12] The benign prognosis of the local plaques of morphea or of many similar plaques must be recognized. An ivory or golden hue, with a livid peripheral border, is characteristic. Rarely, these lesions are ulcerated or bullous; in such cases, they may coexist or merge with lichen sclerosus et atrophicus. Treatment with chloroquine has reduced the inflammation in some patients.[13]

Linear scleroderma is found principally in young children. Roentgenologic evidence of vertebral anomalies was present in 47 per cent of a series of such patients seen at the Clinic. Spina bifida occurred in 20 of 68 patients who had roentgenologic studies. Disorders of pigmentation, such as vitiligo and café au lait spots, are common. In addition, epithelial nevi may be associated with the sclerosis. The classic type of linear scleroderma is seen in Fig. 5, in which the appearance of a "coup de sabre" (saber blow) is associated with facial hemiatrophy. Hemiatrophic states may be present at birth or may appear later in life. In a series of 132 such patients recently reviewed,[14] most of the findings were well established by the twentieth year. The most frequent association with hemiatrophic states is localized scleroderma, principally linear scleroderma; 60 of the

132 patients had such an association. Hemiatrophy does not involve just the skin but also may include fatty tissues and, at times, cartilage and bone. It occurs predominantly in young women, particularly those with a neuropathic heredity. Multiple plastic surgical procedures can be useful in restoring the cosmetic appearance of the face. Occasionally, the surgical creation of artificial arteriovenous communications may help the growth of a limb.

Fig. 5.—The "coup de sabre" (saber blow) lesion of linear scleroderma was associated with hemiatrophy in this patient.

Generalized scleroderma occurs in two forms, namely, that associated with Raynaud's phenomenon, which usually begins in the acral portions of the body, and diffuse scleroderma, which is not associated with vascular changes and has a central distribution. Patients with diffuse scleroderma have a poor prognosis; death usually occurs in 2 years. Acrosclerosis may be rapidly progressive, as recently demonstrated,[15] or it may follow a slow course. Raynaud's phenomenon may precede the visceral and cutaneous sclerosis by as much as 20 years. A recent study of calcinosis cutis in scleroderma emphasized the difference between acrosclerosis and diffuse scleroderma.[16] Calcinosis cutis occurred only in patients with acrosclerosis, usually appearing after 10 years of the disease. Patients who had diffuse scleroderma died too quickly to show the deposition of calcium in the skin. Some patients with acrosclerosis have telangiectasia of the middle portion of the face and trunk, and occasionally they display nodular angiomatous lesions of

the mucous membranes. These lesions occur in long-standing acrosclerosis and must be differentiated from the changes in Rendu-Osler-Weber disease.

Involvement of the gastrointestinal tract, particularly the esophagus, is an extremely useful diagnostic point in cutaneous sclerosis. Studies of esophageal motility extend the knowledge that may be obtained by means of roentgenographic studies made after ingestion of barium. Values for serum carotene, assays of fecal fat, and roentgenologic examination of the small intestine for deficiency patterns are also important. The widening of the periodontal membrane in patients with acrosclerosis who still possess their teeth is useful diagnostically. Measurement of pulmonary function in patients with diffuse linear fibrosis of the lungs is helpful. The electrocardiogram does not reveal any specific abnormality. Cardiac and renal lesions occur late in the disease and have features that are distinctive pathologically but not clinically.

The treatment of progressive systemic sclerosis is still unsatisfactory. No known method of treatment will delay the course of rapidly progressive disease. Chelation accomplished by the intravenous use of edathamil disodium (EDTA) has proved useful, for it reduces the edema and increases the mobility of the involved parts in patients who have acrosclerosis.[17,18] Its success in the childhood forms is most noteworthy. The action of this agent does not depend on extraction of calcium from tissue, but it has a profound effect on the urinary excretion of amino acids. Control of the vascular changes by the administration of phenoxybenzamine hydrochloride (Dibenzyline) is important. A dose of 10 mg. given three to six times daily is helpful when the treatment is tolerated. Physical therapy and a warm environment are necessary to the maintenance of progress.

COMMENT AND CONCLUSIONS

The cutaneous findings in lupus erythematosus, dermatomyositis, and scleroderma are distinctive features and aid in the clinical diagnosis. However, Raynaud's phenomenon may occur in all three instances, and sclerosis has been associated with each disease. Thus, it is the pattern of the individual lesions and their correlation with the course of the disease that are of diagnostic usefulness. In each instance, clinical and laboratory evaluation will indicate the true nature of the primary process.

The treatment of these three conditions is entirely different. Antimalarial compounds are of no value in the treatment of dermatomyositis or scleroderma of the generalized type. Steroids, which help the inflammatory changes in dermatomyositis and systemic lupus erythematosus, do not materially affect any phase of scleroderma. To date, chelation appears to be useful only in the slowly progressive or edematous acrosclerotic form of scleroderma.

While the causes of these three conditions are not known, it is apparent that the basic mechanisms of these diseases are different and that the tissue reactions are widely separated, judging from bedside and laboratory evidence. Close attention to the cutaneous findings in these conditions may allow earlier recognition of the primary nature of the disease, so that appropriate treatment may be given.

REFERENCES

1. Montgomery, Hamilton, and McCreight, W. G.: Disseminate Lupus Erythematosus, Arch. Dermat. & Syph. **60**:356, 1949.
2. Harvey, A. M., Shulman, L. E., Tumulty, P. A., Conley, C. L., and Schoenrich, Edyth H.: Systemic Lupus Erythematosus: Review of the Literature and Clinical Analysis of 138 Cases, Medicine **33**:291, 1954.
3. Ormsby, O. S., and Montgomery, Hamilton: Diseases of the Skin, ed. 8, Philadelphia, 1954, Lea & Febiger.
4. Bartholomew, L. G., Hagedorn, A. B., Cain, J. C., and Baggenstoss, A. H.: Hepatitis and Cirrhosis in Women With Positive Clot Tests for Lupus Erythematosus, New England J. Med. **259**:947, 1958.
5. Muehrcke, R. C., Kark, R. M., Pirani, C. L., and Pollak, V. E.: Lupus Nephritis: A Clinical and Pathologic Study Based on Renal Biopsies, Medicine **36**:1, 1957.
6. Muller, S. A., Winkelmann, R. K., and Brunsting, L. A.: Calcinosis in Dermatomyositis: Observations on Course of Disease in Children and Adults, A.M.A. Arch. Dermat. **79**:669, 1959.
7. Brunsting, L. A.: Dermatomyositis and Systemic Lupus Erythematosus in Childhood, J. Oklahoma M. A. **52**:9, 1959.
8. Curtis, A. C., Blaylock, H. C., and Harrell, E. R., Jr.: Malignant Lesions Associated With Dermatomyositis, J.A.M.A. **150**:844, 1952.
9. Christianson, H. B., Brunsting, L. A., and Perry, H. O.: Dermatomyositis: Unusual Features, Complications and Treatment, A.M.A. Arch. Dermat. **74**:581, 1956.
10. Winkelmann, R. K.: Dermatomyositis, *in* Cohn, H. F., editor: Current Therapy, ed. 12, Philadelphia, 1960, W. B. Saunders Company, pp. 462-463.
11. de Moragas, J. M., Perry, H. O., and Fleisher, G. A.: Serum Glutamic Oxalecetic Transaminase in Dermatomyositis, J.A.M.A. **165**:1936, 1957.
12. Christianson, H. B., Dorsey, C. S., O'Leary, P. A., and Kierland, R. R.: Localized Scleroderma: A Clinical Study of Two Hundred Thirty-Five Cases, A.M.A. Arch. Dermat. **74**:629, 1956.
13. Winkelmann, R. K., and Brunsting, L. A.: The Treatment of Scleroderma, Minnesota Med. **42**:13, 1959.
14. Christianson, H. B.: Hemiatrophy Associated With Scleroderma and Other Conditions. Thesis, Graduate School, University of Minnesota, 1956.
15. Farmer, R. G., Gifford, R. W., Jr., and Hines, E. A., Jr.: Raynaud's Phenomenon as a Guide to Prognosis in Systemic Scleroderma (unpublished data).
16. Muller, S. A., Brunsting, L. A., and Winkelmann, R. K.: Calcinosis Cutis: Its Relationship to Scleroderma, A.M.A. Arch. Dermat. **80**:15, 1959.
17. Winkelmann, R. K.: Treatment of Systemic Scleroderma, Proc. Staff Meet. Mayo Clin. **34**:55, 1959.
18. Muller, S. A., Brunsting, L. A., and Winkelmann, R. K.: The Treatment of Scleroderma With the New Chelating Agent, Edathamil, A.M.A. Arch. Dermat. **80**:187, 1959.

Vol. 196

THE CLINICAL MANIFESTATIONS OF SYSTEMIC LUPUS ERYTHEMATOSUS

A Study of Forty-Five Patients

J. Claude Bennett, M.D.,
James Claybrook, M.D.,
Herbert Kinsey, M.D., and
Howard L. Holley, M.D.

Birmingham, Ala.

From the Division of Rheumatic Diseases, Department of Medicine, Medical College of Alabama, and the Veterans Administration Hospital

(Received for publication Aug. 3, 1960)

DURING the last 10 years, the concept of the clinical course, prognosis, and symptomatology of systemic lupus erythematosus (SLE) has undergone dramatic revision. Prior to this time, it was considered to be a malignantly progressive and invariably fatal disease. Today it is generally recognized that in 20 to 30 per cent of all patients with SLE, the disease undergoes spontaneous remission. Likewise, the use of the LE cell test has made it possible for the clinician to confirm the diagnosis in many early and mild instances of the disease before the full clinical picture becomes apparent. Certainly the introduction of steroid therapy has given a means whereby the severity of clinical symptoms, if not the ultimate outcome of the disease, can be altered.

In this article is presented an analysis of the clinical data obtained from 45 patients studied in the University of Alabama Medical Center. We will also attempt to evaluate the findings in relation to observations reported in other clinical series.

MATERIALS AND METHODS

The diagnosis of SLE has been established in 45 patients seen by the medical or dermatologic service at the University of Alabama Medical Center in the period 1952 through 1959. The records of these patients have been carefully analyzed, and we have had the opportunity to examine and follow up most of them over a period of several years.

In all but 1 patient, positive LE cell tests were obtained. In this patient, a definite diagnosis was made only at autopsy. At no time was a diagnosis of SLE made in the absence of a compatible clinical picture, i.e., on the basis of a positive LE cell test alone.

This work has been supported by a grant from the National Institutes of Health (USPH A-3555) and the John R. Irby Fund for the Study of Arthritis.

The incidence data have been analyzed for the various clinical manifestations of SLE. In general, the time of onset of the disease has been considered to be the time of appearance of symptoms known to be statistically related to SLE. No attempt has been made to compare methods of therapy systematically. Because of the many variables present, it is obvious that this relatively small group of patients does not allow anything approaching an accurate statistical analysis. Nevertheless, it is felt that certain impressions can be gained which will be helpful in understanding the clinical course of this disease.

RESULTS

Race and Sex Distribution.—In the group of patients with an established diagnosis of SLE encountered in this series, there was a male to female distribution of 1:5.4, as can be seen from Table I.

Age.—The range of ages at the time of the initial manifestation was 10 to 59 years. In this group of patients, there seemed to be a definite sex differential for the average age at onset of the disease, 24.9 years for the females and 44 years for the males. No significant difference according to race was noted.

Initial Manifestations.—In most previous clinical reviews, joint symptoms, fever, malaise, and skin manifestations have been the most commonly noted initial manifestations. In general, a similar incidence was noted in this series (Table II).

Establishment of the Diagnosis.—The diagnosis of SLE was supported by a positive LE cell test in 44 of the 45 patients. In the remaining patient, the autopsy findings were compatible with this diagnosis and included periarterial splenic fibrosis and wire-loop lesions in the kidney.

SLE was the original diagnosis made in 26 patients. Chronic discoid lupus was the primary diagnosis in 2 patients. The average time from the initial symptom to the diagnosis of SLE (by no means the same duration as from initial medical examination) was 3.5 years. Rheumatoid arthritis, rheumatic fever, and "nephritis" were frequently encountered as the initial diagnosis in the remaining cases.

Constitutional Symptoms.—The constitutional symptoms demonstrated by this group of patients are given in Table III with the percentage incidence compared with the incidence in other clinical series. Only the history of allergic phenomena stands out as unique to this review, in which approximately one-third of the patients gave such a history. Most of these had a previous history of asthma, childhood eczema, or drug reactions.

Cutaneous Manifestations.—In Table IV are listed the various skin manifestations and their relative incidence as encountered in several reviews. It is interesting that every patient in this study, at some time during the course of illness, had a cutaneous abnormality. The rash occurred over the cheeks and bridge of the nose and neck in only one-third of the patients, but an additional 15 per cent had a similar blotchy, erythematous, scaling rash on other parts of the body. Perhaps the most unusual finding noted in this table was the occurrence of subcutaneous nodules in 15 per cent of the patients. These nodules were firm,

but not indurated, never measured more than 1 cm. in diameter, and were predominantly located over the dorsum of the wrists and the proximal interphalangeal joints of the hands. They were invariably transient, never lasting longer than 48 hours, and often occurred in groups of four or five at a time. Nodules are not unique to this series, as Harvey and colleagues[4] found them in 10 of their 105 patients. Bennett, Zeller, and Bauer[6] found them to be indistinguishable on histologic examination from those found in rheumatoid arthritis or rheumatic fever.

All 4 of the patients in this study with purpura had thrombocytopenia. The purpura was the presenting complaint in only 1 patient.

TABLE I

RACE AND SEX	NUMBER
White males	6
Negro males	1
White females	14
Negro females	24

TABLE II. INITIAL MANIFESTATIONS

SYMPTOMS	INCIDENCE	PERCENTAGE INCIDENCE
Arthralgia or arthritis	31	69
Fever and malaise	6	13
Cutaneous manifestations	5	11
Edema	3	7

TABLE III. CONSTITUTIONAL SYMPTOMS

SYMPTOM	PERCENTAGE INCIDENCE IN SEVERAL SERIES					
	THIS SERIES	SHEARN[1] (34 PATIENTS)	JESSAR[2] (44 PATIENTS)	DUBOIS[3] (62 PATIENTS)	HARVEY[4] (105 PATIENTS	ARMAS-CRUZ[5] (108 PATIENTS)
Fever	93	100	95	97	86	93
Weight loss (15 pounds)	44	74	100	82	71	74
Allergic manifestations	27	—	—	—	—	—
Malaise (extreme)	40	—	—	—	—	68

Musculoskeletal Manifestations.—Joint symptoms were present in all 45 patients, but only 75 per cent had objective signs of arthritis, that is, synovitis, local erythema, and tenderness, or joint deformity, on physical examination (Table V). Localized myositis in muscles of the extremities was present in 2 (4 per cent) of these patients. In patients who had objective arthritis, this was the major complaint throughout the course of the disease. Those with migratory

polyarthralgia usually had an associated dermatitis or serositis which was the most constant source of discomfort. No x-ray changes pathognomonic of SLE were found. In general, changes indistinguishable from rheumatoid arthritis, such as narrowing of the joint space and demineralization of bone adjacent to the joint, were the only abnormalities noted.

TABLE IV. CUTANEOUS MANIFESTATIONS

MANIFESTATION	PERCENTAGE INCIDENCE					
	THIS SERIES	SHEARN[1] (34 PATIENTS)	JESSAR[2] (44 PATIENTS)	DUBOIS[3] (62 PATIENTS)	HARVEY[4] (105 PATIENTS)	ARMAS-CRUZ[5] (108 PATIENTS)
Alopecia	20	—	6	51	3	20
Increased pigmentation	22	—	6	16	12	35
Subcutaneous nodules	15	—	—	—	10	—
Butterfly rash	33	—	—	43	39	60
Purpura	9	15	—	—	9	12
Photosensitivity	22	58	—	40	11	—
Angular stomatitis or cheilosis	9	—	—	—	—	—
Erythematous rash other than over malar area	15	—	—	—	—	—
Exfoliating dermatitis	2	—	—	—	—	—
Raynaud's phenomenon	2	—	—	—	7	16

TABLE V. MUSCULOSKELETAL MANIFESTATIONS

MANIFESTATION	PERCENTAGE INCIDENCE					
	THIS SERIES	SHEARN[1] (34 PATIENTS)	JESSAR[2] (44 PATIENTS)	DUBOIS[3] (62 PATIENTS)	HARVEY[4] (105 PATIENTS)	ARMAS-CRUZ[5] (108 PATIENTS)
Arthritis	75	85	77	90	90	91
Arthralgia	25	—	—	—	—	—
Myositis	4	—	—	—	—	—
No musculoskeletal manifestations	0	—	—	—	—	—

Reticuloendothelial System Manifestations.—Hepatomegaly can occur in patients with SLE, but it is seldom an outstanding finding and clinical jaundice is unusual. In various series, the incidence of hepatomegaly varies from 24 to 44 per cent. Twenty-seven per cent of the patients in this study showed definite hepatic enlargement on physical examination (Table VI). Postmortem study in 2 of these patients revealed only moderate fatty degenerative changes and minimal hepatocellular necrosis. This is in agreement with the findings of Griffith and Vural[7] in their studies on autopsy material obtained in 18 cases.

The reported incidence of splenomegaly in SLE has varied widely. In the present series, the incidence was relatively low (7 per cent). It is of interest that

Dubois[3] found enlargement of the spleen in only 8 per cent of his patients, and he believed that when present there were usually associated hepatomegaly and hemolytic anemia. Armas-Cruz and colleagues[5] are of the opinion that the importance of splenomegaly in this disease has been overemphasized and that when present it is usually of quite mild degree. In their experience, marked enlargement of the spleen militates against the diagnosis of SLE.

TABLE VI. RETICULOENDOTHELIAL MANIFESTATIONS

MANIFESTATION	PERCENTAGE INCIDENCE					
	THIS SERIES	SHEARN[1] (34 PATIENTS)	JESSAR[2] (44 PATIENTS)	DUBOIS[3] (62 PATIENTS)	HARVEY[4] (105 PATIENTS)	ARMAS-CRUZ[5] (108 PATIENTS)
Lymph node enlargement	70	68	37	42	54	46
Hepatomegaly (2 to 5 fingerbreadths below the right costal margin)	27	44	29	34	32	24
Splenomegaly	7	41	27	8	15	19

As can be seen from Table VI, significant generalized or localized lymph node enlargement was found in approximately half of all patients studied with systemic lupus. Two of our patients had such striking lymph node enlargement that a diagnosis of lymphoma was originally entertained.

Cephalin flocculation and thymol turbidity elevations were found in 24 and 18 per cent of our patients, respectively. A reversed albumin to globulin ratio was observed in 47 per cent.

Pulmonary Manifestations.—It has been well recognized for some time that pleural effusion and the pleuritic type of chest pains are common in patients with SLE. In our series, a history of pleuritic-type pain was found in 31 per cent of the patients. Typical pleuritic attacks, but without x-ray findings of effusion (dry pleurisy), have been reported by Harvey and associates,[4] but all of our patients with symptoms of pleurisy were found to have some degree of pleural fluid.

It was in the report of Harvey and associates that attention was called to the presence of parenchymatous pulmonary involvement in SLE. For the most part, this is characterized by a chronic interstitial pneumonitis which often progresses to atelectasis. The alveolar walls and perivascular connective tissue are the sites of primary involvement. Clinically, this lesion is manifested by dyspnea, chronic nonproductive cough, and x-ray changes which show linear and nodular infiltrates. These findings are of long duration and are slow to regress except under steroid therapy. The frequent confusion of these findings with bacterial pneumonia is understandable. In our series, the parenchymal lung lesion was found in only 13 per cent, whereas in Harvey's series it was recognized in 45 per cent of the patients. Table VII shows the comparative incidences of major pulmonary manifestations in several reported clinical series.

Only 2 patients (4 per cent) experienced hemoptysis, and this was minimal. Contrary to expectations, significant hilar lymph node enlargement was found in only 2 patients, being associated with chronic interstitial pneumonitis.

Cardiac Manifestations.—Various cardiac signs and symptoms may be encountered in patients with SLE. As can be seen from Table VIII, 80 per cent of the patients in this series had one or more abnormal cardiac findings. More than

TABLE VII. PULMONARY MANIFESTATIONS

MANIFESTATION	PERCENTAGE INCIDENCE					
	THIS SERIES	SHEARN[1] (34 PATIENTS)	JESSAR[2] (44 PATIENTS)	DUBOIS[3] (62 PATIENTS)	HARVEY[4] (105 PATIENTS)	ARMAS-CRUZ[5] (108 PATIENTS)
Chronic cough	11	—	—	—	—	—
Pleural effusion	31	24	39	55	16	54
Pneumonitis	13	—	20	—	45	Includes all pulmonary manifestations
Pleuritic-type chest pain	31	50	—	60	56	
Hilar adenopathy	4	—	—	—	—	
Hemoptysis	4	—	—	—	—	
No pulmonary manifestations	20	—	—	—	—	

TABLE VIII. CARDIAC MANIFESTATIONS

MANIFESTATION	PERCENTAGE INCIDENCE					
	THIS SERIES	SHEARN[1] (34 PATIENTS)	JESSAR[2] (44 PATIENTS)	DUBOIS[3] (62 PATIENTS)	HARVEY[4] (105 PATIENTS)	ARMAS-CRUZ[5] (108 PATIENTS)
Systolic murmur	51	71	55	42	44	89
Pericarditis	42	18	23	44	45	22
Cardiac enlargement (by x-ray or fluoroscopy)	51	—	34	—	15	—
Auricular fibrillation	4	—	—	—	—	—
Hypertension	11	32	18	—	14	—
Tachycardia	31	—	—	—	—	—
Myocarditis	42	21	—	18	16	—
No cardiac manifestations	20	—	—	—	—	—

half of our patients had systolic murmurs that were usually present at the apex or along the left sternal border. In most instances, these murmurs were associated with cardiomegaly, anemia, or tachycardia. In general, these observations have been noted by other investigators (Table VIII). Even though an apical blowing systolic murmur is heard in many patients with SLE, careful necropsy has revealed little or no correlation between this clinical finding and the nonbacterial verrucous endocarditis of Libman and Sacks.[8] This particular lesion,[9] since its initial correlation with SLE in 1932,[10] has been considered pathognomonic of this disease. Anatomically, it is characterized by soft and friable verrucae located along the line of implantation of valve leaflets and on the mural endocardium.

Fever of significant degree is usually accompanied by tachycardia, but it is of note that the heart rate may be quite rapid in SLE even when temperature is normal and when anemia and cardiac failure are absent. Myocarditis, characterized by cardiac enlargement, gallop rhythm, and electrocardiographic changes, was observed at some time in the clinical course of 42 per cent of our patients. Clinical evidence of this abnormality was often of short duration and was usually not associated with signs of overt congestive heart failure. However, pericarditis characterized by a friction rub and effusion was commonly found in association with evidence of myocarditis. In no instance was the pericarditis associated with suppuration or calcification. Abnormalities of rhythm, other than sinus tachycardia, were found in only 2 patients. Both of these patients had auricular fibrillation associated with clinical evidence of myocarditis.

TABLE IX. GASTROINTESTINAL MANIFESTATIONS

MANIFESTATION	PERCENTAGE INCIDENCE					
	THIS SERIES	SHEARN[1] (34 PATIENTS)	JESSAR[2] (44 PATIENTS)	DUBOIS[3] (62 PATIENTS)	HARVEY[4] (105 PATIENTS)	ARMAS-CRUZ[5] (108 PATIENTS)
Nausea and vomiting	27	77	18	76	25	17
Diarrhea	20	—	—	21	8	—
Melena	11	—	—	—	5	—
Abdominal pain	35	35	22	37	10	—
Ascites	13	15	—	24	—	—
Dysphagia	4.5	—	—	—	6	1

Hypertension has been noted as a significant cardiovascular abnormality in SLE by several groups of investigators (Table VIII). It is not clear whether this is a direct result of the vasculitis[11] occurring in SLE or is caused rather by intercurrent complications, such as chronic pyelonephritis. Five of our patients were observed to have persistent hypertension, but only one of these had clinical evidence of "lupus nephritis." Renal biopsies were not performed in any of these individuals.

Gastrointestinal Manifestations.—Very little is known about the pathologic condition of the gastrointestinal tract in SLE. Certainly, many of these patients have nonspecific abdominal pain, nausea, and vomiting. In some instances, there are undoubtedly inflammatory lesions of the peritoneum, whereas in others mesenteric vasculitis may play the predominant role in the production of these symptoms. The studies of Reifenstein, Reifenstein, and Reifenstein[12] tend to favor peritonitis as the cause of the gastrointestinal symptoms because it was found in 72 per cent of their patients on whom autopsy was performed.

In our series, approximately one-fifth of the patients had nausea and vomiting. This is considerably less than the 75 per cent incidence reported by Dubois[3] and Shearn.[1] The incidence reported by Harvey is more in agreement with our observations. Melena was noted in 5 of our patients (Table IX), but in 3 of these it occurred while they were on steroid therapy.

Abdominal cramping pain, usually periumbilical, was the gastrointestinal symptom (35 per cent) which we found most commonly present. We have seen no incidence of SLE with abdominal pain of such degree as to simulate that experienced during an acute abdominal crisis, as has been occasionally reported.

Ascites occurred in 6 of our patients. All of these had hepatomegaly, together with abnormal liver function tests. Other clinical signs of portal hypertension were not encountered.

TABLE X. GENITOURINARY MANIFESTATIONS

MANIFESTATION	PERCENTAGE INCIDENCE—PRESENT SERIES
Pyuria	53
Hematuria	51
Proteinuria	51
Nocturia, dysuria, and frequency	38
Casts in urine sediment	56
Azotemia	22
Negative urinalysis	15
Menstrual irregularities (amenorrhea, metrorrhagia, menorrhagia)	38

MANIFESTATION	PERCENTAGE INCIDENCE—OTHER SERIES				
	SHEARN[1] (34 PATIENTS)	JESSAR[2] (44 PATIENTS)	DUBOIS[3] (62 PATIENTS)	HARVEY[4] (105 PATIENTS)	ARMAS-CRUZ[5] (108 PATIENTS)
Some abnormality such as proteinuria, abnormal sediment, etc.	62	70	57	65	77
Menorrhagia	—	—	—	14	—

Genitourinary Manifestations.—The kidneys are frequently attacked by the pathologic process in SLE. Most patients have some proteinuria, microscopic hematuria, or abnormal urinary sediment. In our series, slightly over half the patients had one or more of these abnormalities. The incidence of similar observations rises as high as 75 per cent (see Table X) in some series.[2,5] Only 7 (15 per cent) of our patients had repeatedly normal urinalyses.

Azotemia with the blood urea nitrogen above 25 mg. per 100 ml. was present in 10 patients, yet the course of the disease was quite prolonged. It is interesting that there seemed to be little tendency for the development of marked renal insufficiency, as only 3 of our patients died of progressive uremia. Although the course of nephritis in SLE may be acute and fulminating, the onset was gradual in our patients and progressed slowly over 1 to 5 years.

Alterations in the glomerular capillaries were first observed by Baehr, Klemperer, and Schifrin.[13] Microscopically, there is endothelial proliferation in the glomerular capillaries, which may sometimes lead to occlusion of the vessel. Hyaline thrombi may be found in association with focal necrosis of the capillary tufts. This may at times give rise to the morphologically designated "wire-loop"

lesion. This consists of an irregularly thickened basement membrane, and it may involve only a single capillary loop or a number of such lesions may occur. Tubular lesions, as shown by Baggenstoss,[14] may be found in severe renal failure in SLE and consist of hydropic vacuolization and fatty change in the proximal convoluted tubules. Interstitial collections of lymphocytes and plasma cells are sometimes seen.

In their classic review of lupus nephritis, Muehrcke and co-workers[15] state that renal failure is the principal cause of death in SLE. More than 75 per cent of patients with SLE show histologic evidence of nephritis at autopsy.

As can be seen in Table X, over half of our patients had persistent pyuria. Ten of these patients also demonstrated bacteria on routine Gram stain of the urine sediment. Therefore, the relative roles of chronic pyelonephritis and lupus nephritis in producing renal changes in these patients is difficult to evaluate.

Seventeen of our 25 premenopausal women complained of menstrual irregularities. No typical pattern of irregularity could be established, but was most often amenorrhea or oligomenorrhea. Menorrhagia was present in 3 patients and was not found to be associated with any demonstrable blood clotting abnormality. Steroid therapy did not seem to enhance menstrual difficulty but rather tended to produce a more normal pattern if a clinical remission of the underlying disease was induced.

ol. 13
1961

TABLE XI. CENTRAL NERVOUS SYSTEM MAINFESTATIONS

MANIFESTATION	PERCENTAGE INCIDENCE					
	THIS SERIES	SHEARN[1] (34 PATIENTS)	JESSAR[2] (44 PATIENTS)	DUBOIS[3] (62 PATIENTS)	HARVEY[4] (105 PATIENTS)	ARMAS-CRUZ[5] (108 PATIENTS)
Psychosis and/or mental aberrations	18	—	9	29	19	—
Peripheral neuropathy	11	—	—	—	—	19
Convulsions	18	15	7	31	17	8
Coma	13	—	—	—	—	—
No central nervous system manifestations	54	—	—	—	—	—

Nervous System Manifestations.—Neurological and mental disturbances in SLE may take a variety of forms. Convulsions, psychoses,[16] peripheral neuropathies,[17] and delirium may be seen in this disease (Table XI). The major nervous system manifestations in this series were psychoses and convulsions. The incidence of such manifestations was quite high in the series of Dubois,[3] being approximately 30 per cent in each category. In other series,[2,5] however, the incidence is less than 10 per cent. Peripheral neuropathies of motor, sensory, or both modalities have been noted in almost one-fifth of the patients reported by Armas-Cruz and co-workers.[5] Similar nerve involvement was observed in 5 of our patients, and it is of interest that it was seen most frequently in association with severe arthritis.

Cerebral dysfunction, whether psychosis, convulsions, or coma, was invariably associated with an acute exacerbation of the disease. This is probably due to a diffuse or focal cerebral vasculitis. The clinical response of patients with acute neurological symptoms to steroid therapy may be dramatic. On the other hand, these symptoms may first occur while the patient is on maintenance therapy with corticosteroids. In such a situation, the patient must be evaluated in terms of the possibility that the central nervous system aberrations are the result of therapy. Two of our patients required both steroid and anticonvulsant therapy for maintenance of a symptom-free clinical state. In his review of central nervous system manifestations of SLE, Glaser[18] points out that the vascular lesions have a predilection for the gray matter of the cerebral cortex, particularly the presulcal areas. This might explain the frequent occurrence of convulsive seizures. At autopsy, hemorrhage and encephalomalacia, secondary to vascular damage, may often be found in these areas. Forty-six per cent of the patients in our series showed signs and symptoms of neuropsychiatric disease.

Ocular Manifestations.—The ophthalmologic findings in SLE were described some years ago by Maumenee[19] and consist of five types of lesions: (1) Small yellowish white spots called cytoid bodies, located in superficial retinal layers, and similar in appearance to the cotton-wool exudates of hypertensive retinitis; these probably represent degeneration of the nerve fiber layer of the retina. (2) Small hemorrhages in the superficial nerve fiber layer of the retina. (3) Slight papilledema in the form of peripapillary edema,[8] which on rare occasions produces measurable disc elevation. (4) Subretinal edema manifested as a blurring of the choroidal reflex. (5) Histologic evidence of choroiditis with a slight round cell infiltration of the choroid layer. In their review of ocular manifestations in SLE, Hollenhorst and Henderson[20] also noted, in addition to the above abnormalities, embolic petechiae, arterial occlusion, and perivasculitis.

Cotton-wool exudates and uveitis were each seen in 3 of our patients. This is a somewhat lower incidence than that reported in other series (Table XII). Overt papilledema was seen in only 1 of our patients, and this individual was in terminal uremic coma. Jarcho[8] believes that these funduscopic changes are more apt to appear in SLE patients who have signs of central nervous system involvement. This does not seem to be the case except in so far as they correlate with the severity of the disease process and its widespread systemic involvement.

Among our patients, the most commonly observed ocular abnormality was conjunctival and periorbital edema, an incidence of 11 per cent (Table XII). This is a manifestation which has recently been emphasized by Armas-Cruz and colleagues,[5] occurring in 20 per cent of 108 patients. They feel that it is a sign of considerable diagnostic importance.

Hematologic Manifestations.—Alteration of the cellular constituents of the blood has long been recognized as an integral part of the disease. The majority of patients have some degree of anemia, and this factor may indeed be responsible for some of the symptoms which stimulate the patient to consult the physician. The incidence of significant anemia ranges from 60 per cent, as reported by Armas-Cruz and colleagues,[5] to as high as 97 per cent in the series of Shearn and Pirofsky.[1] Three pathogenic mechanisms have been implicated for production of anemia

in SLE.[21] The first type is that generally associated with severe systemic disease and is characterized by being normocytic, with a normal or diminished number of reticulocytes. This relative bone marrow depression is thought to be due largely to disturbed protein metabolism and consequently decreased hemoglobin synthesis. Second, azotemia may cause bone marrow depression and thus anemia in patients with SLE. This phenomenon would of course be of importance only when kidney involvement has progressed to the stage of renal insufficiency. Third, a hemolytic anemia that is due to circulating globulin factors may be present. In this latter group, it is usually possible to demonstrate a positive reaction to the Coombs test.[22,23]

Table XIII shows the incidence of anemia in our patients with SLE. No effort has been made to classify the role played by each of the above mechanisms in these patients. The Coombs test was performed in only 17 of our patients and was found to be positive in 2 individuals (Table XIV).

A very high incidence of elevation of the sedimentation rate was found among our patients (Table XIII), as has also been noted in all other series where this observation has been recorded. This phenomenon, which is believed to be due to abnormalities of serum protein and increased fibrinogen, would be expected to occur with great frequency in a disease which is characterized by flagrant immunoprotein disturbances.

Leukopenia, primarily involving the lymphocytic series, was present in 49 per cent of our patients (Table XIII). Approximately the same incidence has been found in other series. Experimental proof for the mechanism of white cell depression in SLE is still lacking. There is some evidence, however, that antileukocytic "antibodies" may be present in the serum.[24] Leukopenia in a febrile patient may be a useful diagnostic clue. Severe depression with total white cell counts of less than 2,000 is rarely seen in SLE. Mild leukocytosis (10,000 to 15,000) was seen in 8 of our patients and was not related to infection, severity of disease, or steroid therapy.

Moderate thrombocytopenia may be seen in as many as 50 per cent of SLE patients.[21] Rabinowitz and Dameshek[25] have accumulated a series of patients in whom the presenting manifestations were those of thrombocytopenic purpura, although such a severe platelet deficiency is not commonly seen in the over-all picture of SLE. The incidence of platelet deficiency in our series was 35 per cent of 17 patients on whom absolute platelet counts were performed (Table XIII). In only 2 cases was purpura present. It is probable that antiplatelet factors are present in the serum of SLE patients,[26] but proof of their role in causing thrombocytopenia in vivo has yet to be established.

Immunologic Manifestations.—Almost every patient with SLE will have one or more abnormal serologic reactions. Many will have an associated hyperglobulinemia as was found in over half of our patients (Table XIV). The observation of Haserick, Bortz, and Lewis[27] that a specific gamma globulin is involved in the formation of LE cells has led to many studies which indicate that an antigen-antibody type of reaction occurs with this phenomenon.[28-30]

The common occurrence of false positive serologic reactions for syphilis in patients with SLE was first noted by Coburn and Moore.[31] The incidence ranges

TABLE XII. EYE MANIFESTATIONS

MANIFESTATION	PERCENTAGE INCIDENCE					
	THIS SERIES	SHEARN[1] (34 PATIENTS)	JESSAR[2] (44 PATIENTS)	DUBOIS[3] (62 PATIENTS)	HARVEY[4] (105 PATIENTS)	ARMAS-CRUZ[5] (108 PATIENTS)
Conjunctival and periorbital edema	11	—	—	—	—	20
Conjunctivitis and keratoconjunctivitis	4	—	—	—	5	—
Visual blurring with diplopia	7	—	—	—	—	—
Cotton-wool exudates in fundi (cytoid bodies)	7	—	14	—	24	17
Papilledema	2	—	—	—	2	2
Uveitis	7	—	—	—	—	—
Hemorrhages	0	—	11	—	10	3
No eye manifestations	67	—	—	—	—	—

TABLE XIII. HEMATOLOGIC MANIFESTATIONS IN PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS

MANIFESTATION	NUMBER STUDIED	NUMBER ABNORMAL	PERCENTAGE ABNORMAL
Anemia (Hb. less than 11 Gm. per 100 ml. or Hct. less than 35)	45	33	73
Elevated sedimentation rate (Above 20 Wintrobe)	41	40	97
Leukopenia (Less than WBC 4,000)	45	22	49
Leukocytosis (More than WBC 10,000)	45	8	15
Elevated absolute eosinophil count (Above 350 per cubic millimeter)	24	0	0
Thrombocytopenia (Less than 250,000 per cubic meter)	17	6	35

TABLE XIV. IMMUNOLOGIC MANIFESTATIONS IN PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS

TEST	NUMBER STUDIED	NUMBER POSITIVE	PERCENTAGE POSITIVE
Hyperglobulinemia (Above 3.5 Gm. per 100 ml.)	40	25	62
LE cell preparation	45	44	98
Serologic test for syphilis	37	5	13
Latex fixation	17	4	24
Coombs test	17	2	12
SLE skin test	17	15	89
DNA precipitation	12	12	100
Aggregated gamma globulin–precipitation	12	6	50

from 9 to 87 per cent, according to various authors. In a follow-up study of 148 instances of the false reactors, Moore and Lutz[32] found that 10 per cent subsequently developed SLE and 45 per cent had evidence of a "collagen vascular disease." The incidence of positive reactions with cardiolipin in our series was 10 per cent.

As can be seen from Table XIV, a significant proportion of our patients showed positive latex fixation reactions and precipitation reactions with aggregated gamma globulin. Whether these positive tests, which are more commonly seen in rheumatoid arthritis, are due to the presence of the "rheumatoid factor" is not yet clear. These findings again emphasize interrelationships between SLE and rheumatoid arthritis, but their etiologic significance remains to be elucidated.

Cutaneous hypersensitivity to leukocytes[33,34] and positive response to the desoxyribonucleic acid precipitation test[35] are reactions which are more specific for SLE. These tests have been performed in some of our more recently acquired patients, and the results are recorded in Table XIV.

ol. 13
1961

COURSE AND OUTCOME

During the natural history of SLE, one observes periods of exacerbation and remission. The course may be prolonged over many years with gradual involvement of various organ systems to produce the "multisystemic" clinical picture that has been previously discussed. Often the patient may be completely symptom-free for weeks or months without need for therapy. In a few instances, however, the disease begins abruptly and follows a fulminant, unremitting course to death of the patient in a matter of weeks.

Spontaneous remissions are said to occur in one-fourth to one-third of patients with SLE.[3,4] Frequently, a clinical remission is of transient nature and may be present for only several weeks. This can be misleading, however, since a superficial remission may be associated with minor malaise, anorexia, and arthralgia. Minimal abnormalities of the urine such as proteinuria can often persist despite the absence of symptoms of systemic disease. Because of difficulty in deciding what constitutes a clinical "remission," for the purposes of this report, we have arbitrarily defined it as a period of at least 6 months in which the patient (1) is symptom-free, (2) has no evidence of active disease on examination, and (3) has completely normal response to laboratory studies. This strict definition has led to a rather low incidence of remissions in our series. Only 3 patients have continued in complete remission after the initial acute manifestations of the disease. Six additional patients have shown intermittent disease activity, but at the present time only 2 of these have a sustained remission of more than 6 months' duration.

From our small series, it is impossible to draw any conclusions regarding therapy. Nevertheless, one does gain the impression that steroid therapy tends to minimize the manifestations of disease activity even if the ultimate outcome is unaffected. Early reports of therapy with steroids and ACTH in SLE gave enthusiasm for their therapeutic benefit.[36-38] Although this initial exuberance has

waned, it is still felt by most workers that a significant percentage of patients with SLE can be rehabilitated by judicious steroid therapy.[39,40] In most instances, a reasonably good balance of health can be maintained with continuous or intermittent treatment. Two of our patients with acute SLE promptly went into remission after steroid therapy was instituted; both have now been symptom-free for 4 years. Such good results are not the rule, however, and most patients have to be on long-term maintenance therapy with steroids. In accordance with the recent work of Ziff, Esserman, and McEwen,[41] it would appear that a "combination" therapeutic program using antimalarials can reduce the maintenance dose of steroids.

Adequately controlled studies to determine the therapeutic efficacy of steroids in the more malignant manifestations of SLE such as nephritis and endocarditis have not been reported. Such a study would give us more information as to whether steroids actually alter the outcome or only suppress the clinical manifestations of this disease.

SUMMARY

The clinical manifestations, course, and outcome of a group of 45 patients with systemic lupus erythematosus have been presented. It is hoped that through recognition of this disorder early in its course and appreciation of therapeutic alteration of the disease picture, a better understanding of this condition can be reached.

REFERENCES

1. Shearn, M. A., and Pirofsky, B.: Disseminated Lupus Erythematosus; Analysis of 34 Cases, A.M.A. Arch. Int. Med. **90**:790, 1952.
2. Jessar, R. A., Lamont-Havers, R. W., and Ragan, C.: Natural History of Lupus Erythematosus Disseminatus, Ann. Int. Med. **38**:717, 1953.
3. Dubois, E. L.: The Effect of the L.E. Cell Test on the Clinical Picture of Systemic Lupus Erythematosus, Ann. Int. Med. **38**:1265, 1953.
4. Harvey, A. M., et al.: Systemic Lupus Erythematosus: Review of the Literature and Clinical Analysis of 138 Cases, Medicine **33**:291, 1954.
5. Armas-Cruz, R., Harnecker, J., Ducach, G., Jalil, J., and Gonzaliz, F.: Clinical Diagnosis of Systemic Lupus Erythematosus, Am. J. Med. **25**:409, 1958.
6. Bennett, G. A., Zeller, J. W., and Bauer, W.: Subcutaneous Nodules of Rheumatoid Arthritis and Rheumatic Fever; Pathologic Study, Arch. Path. **30**:70, 1940.
7. Griffith, G. C., and Vural, I. L.: Acute and Subacute Disseminated Lupus Erythematosus; Correlation of Clinical and Postmortem Findings in 18 Cases, Circulation **3**:492, 1951.
8. Jarcho, S.: The Clinical Features of Systemic Lupus Erythematosus, J. Mt. Sinai Hosp. New York **26**:278, 1959.
9. Libman, E., and Sacks, B.: A Hitherto Undescribed Form of Valvular and Mural Endocarditis, Arch. Int. Med. **33**:701, 1924.
10. Gross, L.: The Heart in Atypical Verrucous Endocarditis (Libman-Sacks), *in* Contributions to the Medical Sciences in Honor of Dr. Emanuel Libman by His Pupils, Friends, and Colleagues, New York, 1932, The International Press, vol. 2, p. 527.
11. Sutnick, A. I., Sellers, A. M., Jeffers, W. A., and Wolferth, C. C.: Systemic Lupus Erythematosus and Hypertension Without Renal Insufficiency, Ann. Int. Med. **52**:849, 1960.
12. Reifenstein, E. C., Reifenstein, E. C., Jr., and Reifenstein, G. H.: Variable Symptom Complex of Undetermined Etiology With Fatal Termination, Including Conditions Described as Visceral Erythema Group (Osler), Disseminated Lupus Erythematosus, Atypical Verrucous Endocarditis (Libman-Sacks), Fever of Unknown Origin (Christian) and Diffuse Peripheral Vascular Disease (Baehr and Others), Arch. Int. Med. **63**:553, 1939.
13. Baehr, G., Klemperer, P., and Schifrin, A.: A Diffuse Disease of the Peripheral Circulation (Usually Associated With Lupus Erythematosus and Endocarditis), Tr. A. Am. Physicians **50**:139, 1935.

14. Baggenstoss, A. H.: Symposium on Systemic Lupus Erythematosus; Visceral Lesions in Disseminated Lupus Erythematosus, Proc. Staff Meet. Mayo Clin. **27**:412, 1952.
15. Muehrcke, R. C., Kark, R. M., Pirani, C. L., and Pollak, V. E.: Lupus Nephritis: A Clinical and Pathologic Study Based on Renal Biopsies, Medicine **36**:1, 1957.
16. O'Connor, J. F.: Psychoses Associated With Systemic Lupus Erythematosus, Ann. Int. Med. **51**:526, 1959.
17. Heptinstall, R. H., and Sowry, G. S. C.: Peripheral Neuritis in Systemic Lupus Erythematosus, Brit. M. J. **1**:525, 1952.
18. Glaser, G. H.: Lesions of the Central Nervous System in Disseminated Lupus Erythematosus, A.M.A. Arch. Neurol. & Psychiat. **67**:745, 1952.
19. Maumenee, A. E.: Retinal Lesions in Lupus Erythematosus, Am. J. Ophth. **23**:971, 1940.
20. Hollenhorst, R. W., and Henderson, J. W.: The Ocular Manifestations of the Diffuse Collagen Diseases, Am. J. M. Sc. **221**:211, 1951.
21. Lee, S. L., and Davis, B. J.: The Blood in Systemic Lupus Erythematosus, J. Mt. Sinai Hosp. New York **26**:261, 1959.
22. Michael, S. R., Vural I. L., Bassen, F. A., and Schaefer, L.: Hematologic Aspects of Disseminated (Systemic) Lupus Erythematosus, Blood **6**:1059, 1951.
23. Hargroves, M. M., and Opfell, R. W.: Systemic Lupus Erythematosus and the Blood, Prog. Hematol. **1**:249, 1956.
24. Van Loghem, J. J., Jr., Van der Hart, M., Hijmans, W., and Schuit, H. R. E.: The Incidence and Significance of Complete and Incomplete White Cell Antibodies With Special Reference to the Use of the Coombs Consumption Test, Vox Sang. **3**:203, 1958.
25. Rabinowitz, Y., and Dameshek, W.: Systemic Lupus Erythematosus After "Idiopathic" Thrombocytopenic Purpura: A Review, Ann. Int. Med. **52**:1, 1960.
26. Harrington, W. J., et al.: Immunologic Mechanisms in Idiopathic and Neonatal Thrombocytopenic Purpura, Ann. Int. Med. **38**:433, 1953.
27. Haserick, J. R., Lewis, L. A., and Bortz, D. W.: Blood Factor in Acute Disseminated Lupus Erythematosus; Determination of Gamma Globulin as Specific Plasma Fraction, Am. J. M. Sc. **219**:660, 1950.
28. Godman, G. C., and Deitch, A. D.: A Cytochemical Study of the L.E. Bodies of Systemic Lupus Erythematosus. I. Nucleic Acids, J. Exper. Med. **106**:575, 1957.
29. Godman, G. C., and Deitch, A. D.: A Cytochemical Study of the L.E. Bodies of Systemic Lupus Erythematosus. II. Proteins, J. Exper. Med. **106**:593, 1957.
30. Friou, G. J.: Identification of the Nuclear Component of the Interaction of Lupus Erythematosus Globulin and Nuclei, J. Immunol. **80**:476, 1958.
31. Colburn, A. F., and Moore, D. H.: Plasma Proteins in Disseminated Lupus Erythematosus, Bull. Johns Hopkins Hosp. **73**:196, 1943.
32. Moore, J. E., and Lutz, W. B.: The Natural History of Systemic Lupus Erythematosus: An Approach to Its Study Through Chronic Biologic False Positive Reactors, J. Chron. Dis. **1**:297, 1955.
33. Bennett, J. C., and Holley, H. L.: Intradermal Hypersensitivity in Systemic Lupus Erythematosus. (In press.)
34. Friedman, E. A., Bardawil, W. A., Merrill, J. P., and Hanau, C.: "Delayed" Cutaneous Hypersensitivity to Leukocytes in Disseminated Lupus Erythematosus, New England J. Med. **262**:486, 1960.
35. Deicher, H. R., Holman, H. R. and Kunkel, H. G.: The Precipitin Reaction Between DNA and a Serum Factor in Systemic Lupus Erythematosus, J. Exper. Med. **109**:97, 1959.
36. Grace, A. W., and Combes, F. C.: Remission of Disseminated Lupus Erythematosus Induced by Adrenocorticotropin, Proc. Soc. Exper. Biol. & Med. **72**:563, 1949.
37. Harvey, A. M., et al.: Observations on Effect of Adrenocorticotropic Hormone (ACTH) on Disseminated Lupus Erythematosus, Drug Hypersensitivity Reactions, and Chronic Bronchial Asthma, Tr. Am. Clin. & Climatol. A. **61**:221, 1950.
38. Soffer, L. J., and Bader, R.: Corticotropin and Cortisone in Acute Disseminated Lupus Erythematosus; Results of Long-term Use, J.A.M.A. **149**:1002, 1952.
39. Soffer, L. J.: The Therapy of Systemic Lupus Erythematosus, J. Mt. Sinai Hosp. New York **26**:297, 1959.
40. McCombs, R. P., and Patterson, J. F.: Factors Influencing the Course and Prognosis of Systemic Lupus Erythematosus, New England J. Med. **260**:1195, 1959.
41. Ziff, M., Esserman, P., and McEwen, C.: Observations on the Course and Treatment of Systemic Lupus Erythematosus, Arthritis & Rheumat. **1**:332, 1958.

Cardiovascular

THE FREQUENCY OF ARTERIOSCLEROTIC AND HYPERTENSIVE HEART DISEASE AMONG OSTENSIBLY HEALTHY WORKING POPULATIONS IN FINLAND

An Electrocardiographic and Clinical Study

PENTTI M. RAUTAHARJU, M.D.,*
M. J. KARVONEN, M.D., PH.D.,**
AND ANCEL KEYS, PH.D.***

HELSINKI, FINLAND, AND
MINNEAPOLIS, MINN.

From the Institute of Occupational Health, Helsinki, Finland, and the Laboratory of Physiological Hygiene, University of Minnesota

(Received for publication Sept. 1, 1960)

EPIDEMIOLOGIC studies have shown important differences in the frequency of coronary heart disease (CHD) between populations and among social and activity classes within populations. According to available statistics, the frequency of cardiovascular disease, particularly of CHD, is high in Finland. Moreover, cardiovascular disease appears to be more common in the eastern than in the western part of the country.[1]

These regional mortality and disability data pose the question whether any corresponding differences may be observed among the clinical healthy populations in the same areas. For this purpose, a cooperative investigation was carried out in 1956, which included dietary studies,[2] serum cholesterol and lipoprotein determinations,[1] anthropometric, clinical, and electrocardiographic (ECG) examinations, and the collection of data on smoking[3] of people free from known or manifest disease. Examinations were carried out on two rural groups, and on two population groups examined in Helsinki in 1956, one consisting of firemen of the City of Helsinki and the other of business and professional men. By the same methods, a group of lumberjacks from East Finland was examined by Karvonen and associates[4] in 1959. The frequency of ECG and clinical findings compatible with coronary and hypertensive heart diseases among these five different populations is reported in the present paper.

This study was supported by the Science Board of the State of Finland and by grants in aid to Prof. Ancel Keys from the National Dairy Council (Chicago), the Winton Companies Fund (Minneapolis), and the American Heart Association.

*Research Fellow of the Minnesota Heart Association.

**Director of Physiological Department, Institute of Occupational Health, Helsinki; Docent of Physiology, University of Helsinki, Finland.

***Director, Laboratory of Physiological Hygiene, University of Minnesota, Minneapolis, Minn.

MATERIAL

The study was carried out in Helsinki and in two rural areas. One of them, in the southwest corner of Finland, comprised the following communities: Karinainen, Mellilä, Mietoinen, Nousiainen, Pöytyä, and Vehmaa; the other, in East Finland, Juuka, and Ilomantsi (Fig. 1). In each of these rural regions, a number of villages suitably situated were chosen for the study. With the aid of local medical personnel, all those who were disabled to the extent of not being able to work or known to be ill were excluded from the material. Thereafter, all the men in the age range of 20 to 59 years were visited at their homes by a member of the team and invited to the examination. When necessary, ample time was used for persuasion. Substantially all those invited responded; the only exceptions had legitimate excuses. The two regional groups may be considered as "total ostensibly healthy populations" of the chosen geographic areas.

In Helsinki, two population groups were studied: 270 firemen of the Helsinki Fire Brigade and 90 "directors," professional and business men from the local Rotary and Lions Clubs. The firemen were subjected to the study as a part of their obligatory annual medical examination; hence, the group definitely is a "total population" sample. Active firemen must also be considered as an "ostensibly healthy" group. The group of "directors" was collected from the membership lists of local Rotary and Lions Clubs. In this group, self-selection cannot be excluded though, also here, substantially all those approached responded.

A quite different working population group was studied in East Finland, in Ilomantsi, consisting of 99 lumberjacks.[4] This study was carried out in the logging camps in the forests. All lumberjacks working in the selected area were studied.

The age distribution of the men studied is shown in Table I.

TABLE I. AGE DISTRIBUTION OF THE MEN IN EACH SAMPLE

SAMPLE	AGE (YEARS)				TOTAL
	20–29	30–39	40–49	50–59	
East	69	83	81	49	282
West	68	67	58	51	244
Firemen	63	86	92	29	270
Directors	6	42	31	11	90
Lumberjacks	29	30	24	16	99
Total	235	308	286	156	985

METHODS

The electrocardiograms were recorded with a 2-channel "Mingograf" (Elema Ab). The position of the precordial electrodes differed from the conventional ones in that V_1 and V_2 were recorded at the level of the fifth intercostal space at the sternum, which horizontal level was used in determining the positions for chest leads from V_4R to V_6. This method has the advantage that it reduces

the interindividual variability in determining the positions of electrodes and thus increases the comparability of populations studied. This system has the further advantage that it makes possible the determination of the transition points of QRS and T at the horizontal level for vector analysis.[5]

Classification of ECG Findings.—The ECG criteria and the classification scheme developed and tested in the field studies conducted by the Laboratory of Physiological Hygiene, University of Minnesota, were used.[6] This system appears to be very suitable for application to population surveys and to their subsequent statistical analysis, and provides a scheme for the objective classification of ECG characteristics without entering into the controversial field of subjective diagnostic interpretations. Multiple abnormalities in an individual ECG are treated each as an independent entry.

Diagnostic Entities.—In the classification scheme for the interpopulation comparisons of heart disease suggested by the Laboratory of Physiological Hygiene,[7] a subject is designated to one category only, whereas in the above classification of ECG findings the same subject may occur several times in the statistics. Concerning *coronary heart disease* three categories of diagnosis are provided: (1) myocardial infarction, (2) angina pectoris, and (3) ECG abnormalities in themselves nonspecific, occurring in the absence of hypertension (160 systolic and/or 95 diastolic or over) or other manifest etiology of heart disease.

For *myocardial infarction*, the following criteria were used:

Probable: (1) Specific ECG criteria for probable old infarction (code I, 1); (2) ECG criteria for possible infarction (code I, 2) plus acceptable history of "attack"; (3) verified history of infarction with characteristic acute changes by ECG and/or enzyme test. (Cases meeting this criterion were not included in the present study.)

Possible: (1) ECG criteria for "possible" infarction alone (code I, 2); (2) acceptable history of "attack," verification unavailable, plus nonspecific ECG abnormalities at time of examination (code IV, 1 to VIII, 5: S-T and T wave abnormalities, A-V and intraventricular conduction disturbances, arrhythmias, excluding nodal rhythm, sinus tachycardia, and bradycardia).

For the diagnosis of *angina pectoris*, agreement between examiner and reviewer on history of classical angina pectoris was required.

The blood pressure was measured in the sitting position in association with the routine medical examination. Diastolic pressure was recorded as the Korotkoff sounds finally disappeared. All values given are based on a single measurement only.

Subjects with hypertension (160 systolic and/or 95 diastolic or over) were placed in three categories: (a) hypertension associated with high voltage ECG pattern (code III, 1: R wave > 2.6 mV in V_5, > 2.0 mV I, II, III or aVF, or > 1.2 mV in aVL); (b) hypertension with other abnormalities in ECG (code IV to VIII: S-T and T abnormalities, A-V or intraventricular conduction defects and arrhythmias excluding sinus tachycardia and bradycardia); (c) hypertension, apparently uncomplicated, with no reportable finding of the categories mentioned above.

In the statistical evaluation, the chi-square test was used.

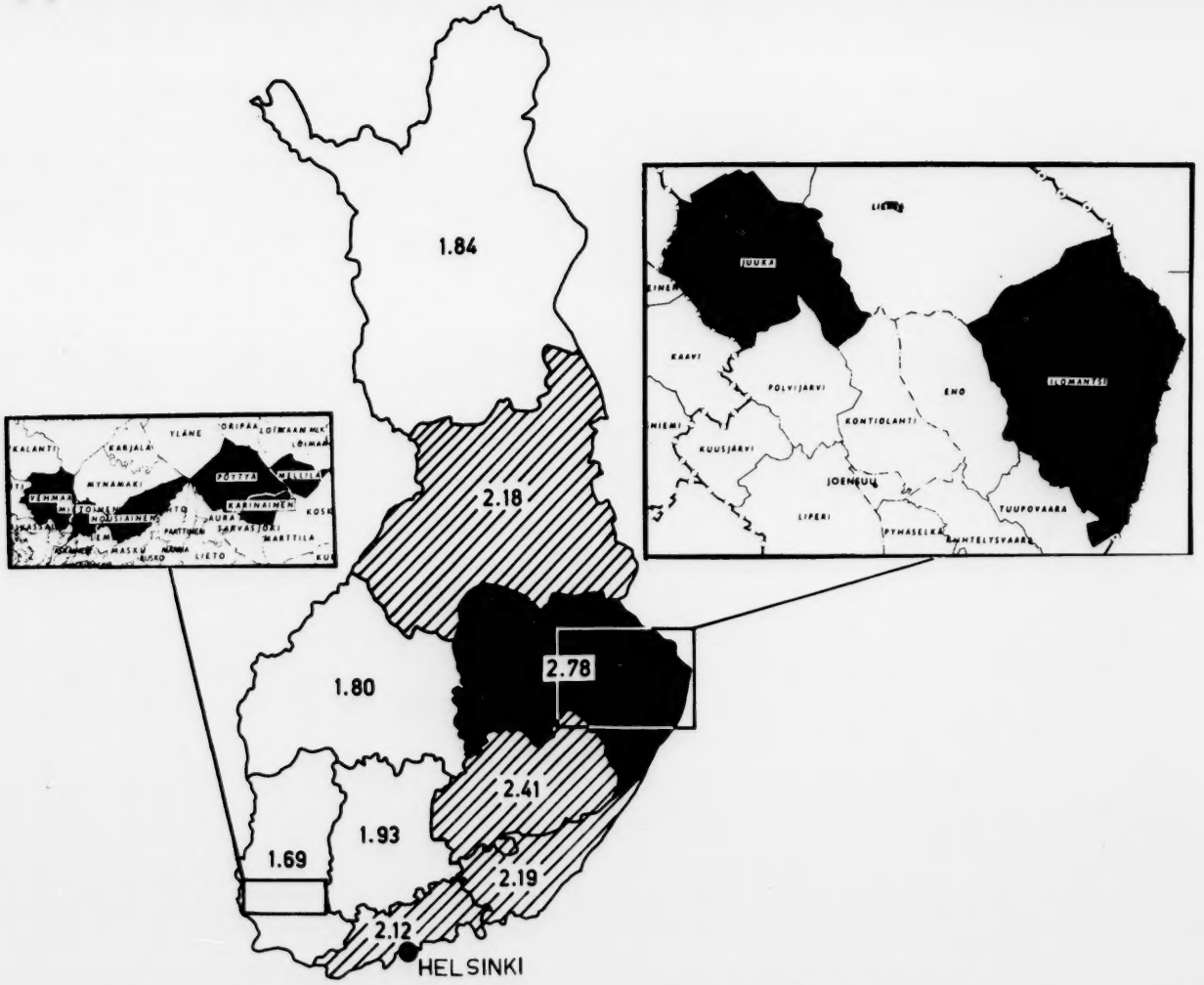

Fig. 1.—A map of Finland showing the communities studied. The numbers on the map indicate mortality per thousand individuals ascribed to arteriosclerotic heart disease (International List, Sixth Revision, Nos. 420-422) in Finland during the years 1951-1955.

RESULTS

Frequency of ECG Abnormalities.—The frequency of ECG abnormalities among different population groups is shown in Table II. It is underlined that the entries in this table refer to ECG patterns, not to individuals. Thus, multiple ECG abnormalities are entered independently. Code numbers correspond to the items of the classification system mentioned above.[6] Axis deviations (II, 1 and 2), sinus tachycardia and sinus bradycardia (VIII, 7, 8) are not included. The total occurrence of findings was 22.4 per 100 ECG's. The frequency was lowest among lumberjacks (13.1 per cent), and at almost the same level among the rural population of West Finland (16.0 per cent) and "directors" (18.9 per cent). The frequency was higher among the rural group of men of East Finland (25.9 per cent) and highest among the firemen of Helsinki Fire Brigade (28.9 per cent). The differences are essentially due to left ventricular preponderance, S-T depressions, and T wave abnormalities.

Frequency of Diagnostic Entities.—The frequency of cases belonging to the *coronary group* is shown in Table III. In this table, each individual is entered only once. The total percentage of cases in this group is 6.0. The frequency is lowest among lumberjacks from East Finland (4.0 per cent) and practically at the same level among the rural men from West Finland (4.1 per cent) and firemen from Helsinki (4.8 per cent). The rate was slightly higher among "directors" from Helsinki (7.8 per cent) and highest among the eastern group of rural men (8.9 per cent). The difference between eastern and western rural groups is significant ($P < 0.05$). It is largely due to angina pectoris and to the category of possible myocardial infarction.

Table IV presents the age distribution of the coronary group.

Table V presents the *hypertensive group*. The frequency of cases interpreted as being hypertensive heart disease is 5.6 per cent. The frequency is highest among firemen (9.3 per cent), and considerably lower among "directors" (3.3 per cent) and the western rural group (3.3 per cent). Hypertensive heart disease tends to be more common in the eastern rural group (6.6 per cent) than in the western one, but the difference does not attain statistical significance. The total frequency of hypertension is significantly higher in the East (37.2 per cent) and among firemen (35.9 per cent) than in the West (25.0 per cent) and among "directors" (25.6 per cent). The frequency among lumberjacks was significantly lower than in any other group (11.4 per cent). (The significance level between East and West was found to the $P < 0.01$, and between "directors" and lumberjacks $P < 0.05$.)

The frequency of high voltage R waves compatible with *left ventricular hypertrophy* (LVH) is shown in Table VI. The total group is subdivided into three subgroups according to the blood pressure, following the recommendation by the World Health Organization[8]: (a) normal pressure, systolic under 140 and diastolic under 90; (b) intermediate elevation of pressure, systolic 140-159 and/or diastolic 90-94; (c) hypertension, systolic 160 or over and/or diastolic 95 or over. One or more of the high voltage criteria were observed in 135 individuals (13.7 per cent), equally distributed in all three blood pressure groups. This was

TABLE II. ELECTROCARDIOGRAPHIC ABNORMALITIES IN DIFFERENT WORKING POPULATION SAMPLES IN THE AGE RANGE 20–59 YEARS: 282 RURAL MALES FROM EAST FINLAND; 244 RURAL MALES FROM WEST FINLAND; 270 FIREMEN FROM FIRE BRIGADE OF HELSINKI; 90 "DIRECTORS" (BUSINESS AND PROFESSIONAL MEN) FROM HELSINKI; 99 LUMBERJACKS FROM EAST FINLAND

ABNORMALITY	CODE NO.	ECG ABNORMALITIES					ABNORMALITIES PER 100 ECG'S					TOTAL
		EAST	WEST	FIRE-MEN	DIREC-TORS	LUMBER-JACKS	EAST	WEST	FIRE-MEN	DIREC-TORS	LUMBER-JACKS	
Probable infarction	I, 1	1	1	1	—	—	0.4	0.4	0.4	0.0	0.0	0.3
Possible infarction	I, 2	4	—	2	—	—	1.4	0.0	0.7	0.0	0.0	0.6
	I, 1–2	5	1	3	—	—	1.8	0.4	1.1	0.0	0.0	0.9
Left ventricular preponderance	III, 1	42	23	54	7	9	14.9	9.4	20.0	7.8	9.1	13.7
Right ventricular preponderance	III, 2	—	1	2	—	—	0.0	0.4	0.7	0.0	0.0	0.3
S-T depression	IV, 1	4	2	3	2	1	1.4	0.8	1.1	2.2	1.0	1.2
	IV, 2	5	2	3	1	1	1.8	0.8	1.1	1.1	1.0	1.2
	IV, 3	4	0	1	—	—	1.4	0.0	0.4	0.0	0.0	0.5
	IV, 1–3	13	4	7	3	2	4.6	1.6	2.6	3.3	2.0	2.9
T wave abnormalities	V, 1	—	—	—	—	—	0.0	0.0	0.0	0.0	0.0	0.0
	V, 2	6	3	1	—	1	2.1	1.2	0.4	0.0	1.0	1.1
	V, 3	5	3	4	4	1	1.8	1.2	1.5	4.4	1.0	1.7
	V, 1–3	11	6	5	4	2	3.9	2.4	1.9	4.4	2.0	2.8
A-V conduction defects (any degree)	VI, 1–4	1	1	3	1	—	0.4	0.4	1.1	1.1	0.0	0.6
Intraventricular conduction defects	VII, 1–4	1	1	2	1	—	0.4	0.4	0.7	1.1	0.0	0.5
Arrhythmias (excluding sinus tachycardia and bradycardia)	VIII, 1–6	—	2	2	1	—	0.0	0.8	0.7	1.1	0.0	0.5
Total No. of abnormalities		73	39	78	17	13	25.9	16.0	28.9	18.9	13.1	22.4

TABLE III. CLASSIFICATION OF THE CASES INTERPRETED AS PROBABLE OR POSSIBLE CORONARY HEART DISEASE INTO THREE MAIN CATEGORIES

CATEGORY	NO. OF CASES					%					
	EAST	WEST	FIRE-MEN	DIREC-TORS	LUMBER-JACKS	EAST	WEST	FIRE-MEN	DIREC-TORS	LUMBER-JACKS	TOTAL
Myocardial infarction											
Probable	1	1	1	0	0	0.4	0.4	0.4	0.0	0.0	0.3
Possible	5*	0	2	0	0	1.8	0.0	0.7	0.0	0.0	0.7
Total	6	1	3	0	0	2.1	0.4	1.1	0.0	0.0	1.0
Angina pectoris	8	1	0	1	0	2.8	0.4	0.0	1.1	0.0	1.0
Isolated ECG abnormalities	11	8	10	6	4	3.9	3.3	3.7	6.7	4.0	4.0
Total	25	10	13	7	4	8.9	4.1	4.8	7.8	4.0	6.0

*Including one man with an acceptable history of "attack," with nonspecific ECG abnormalities.

Vol. 13
1961

TABLE IV. AGE DISTRIBUTION OF CORONARY GROUP IN EACH SAMPLE

SAMPLE	AGE (YEARS)												ALL AGES		
	20–29			30–39			40–49			50–59					
	NO. CASES	TOTAL	%	NO. CASES	TOTAL	%	NO. CASES	TOTAL	%	NO. CASES	TOTAL	%	NO. CASES	TOTAL	%
East	4	69	5.8	4	83	4.8	9	81	11.1	8	49	16.3	25	282	8.9
West	4	68	5.9	2	67	2.9	2	58	3.5	2	51	3.9	10	244	4.1
Firemen	4	63	6.4	2	86	2.3	5	92	5.4	2	29	6.9	13	270	4.8
Directors	3	6	50.0	2	42	4.7	0	31	0	2	11	8.2	7	90	7.8
Lumberjacks	2	29	6.9	1	30	3.3	1	24	4.2	0	16	0	4	99	4.0
Total	17	235	7.2	11	308	3.6	17	286	5.9	14	156	9.0	59	985	6.0

TABLE V. FREQUENCY OF CASES INTERPRETED AS HYPERTENSIVE HEART DISEASE OR UNCOMPLICATED BLOOD PRESSURE ELEVATION

	NO. OF CASES					%					
	EAST	WEST	FIRE-MEN	DIREC-TORS	LUMBER-JACKS	EAST	WEST	FIRE-MEN	DIREC-TORS	LUMBER-JACKS	TOTAL
Hypertension and LVH in ECG (Item III, 1)	16	5	22	1	0	5.7	2.1	8.1	1.1	0.0	4.5
Hypertension with other abnormalities in ECG (Items IV–VIII)	3	3	3	2	0	1.1	1.2	1.1	2.2	0.0	1.1
Total	19	8	25	3	0	6.7	3.3	9.2	3.3	0.0	5.6
Hypertension, apparently uncomplicated	86	53	72	20	11	30.5	21.7	26.7	22.2	11.4	24.6
Total	105	61	97	23	11	37.2	25.0	35.9	25.6	11.4	30.2

Vol. 1
1961

the case in all population groups. Because of the different electrode position in V_1, Sokolov's criteria for left ventricular hypertrophy were not used. It is felt that this sensitive criterion would further increase the number of false positive cases among preselected working population.

TABLE VI. THE FREQUENCY OF HIGH VOLTAGE R WAVE ITEMS COMPATIBLE WITH LEFT VENTRICULAR HYPERTROPHY AS RELATED TO BLOOD PRESSURE

BLOOD PRESSURE	NO. OF SUBJECTS	LVH PATTERN IN ECG	
		NO. OF CASES	%
Normal	421	61	14.5
Intermediate elevation	267	30	11.2
Hypertension	297	44	14.8
Total	985	135	13.7

DISCUSSION

Electrocardiography as a Tool in Epidemiology.—The present study is an attempt to apply electrocardiography to the epidemiologic investigation of so-called healthy populations. A prerequisite for this is that all ECG observations should be strictly defined and standardized. The classification system used provides a tool for objective description of ECG abnormalities without recourse to subjective diagnostic or prognostic interpretation. The diagnostic entities are then based on the combination of these defined ECG observations with physical examination and history.

The coronary group included 17 men in the age range of 20 to 29 years. Possibly many of these diagnoses were "false positives." However, their inclusion does not essentially affect the applicability of this method to population comparisons. The only possibility for finding out which are the "false positives" is to make a long-term follow-up study and to judge the significance of each finding from the subsequent history. Even this approach is not strictly correct, for an ECG abnormality may well be a remnant from an active phase of disease which has later regressed.

The essential point in epidemiology is the recording of objective population characteristics. In a long-term study, statistical information on the prognostic significance of individual signs and symptoms and their combinations may be obtained. However, the diagnosis in any given case may always require scope for arbitrary judgment.

The results of the present study suggest that electrocardiography can successfully be used as an objective tool in the study of cardiovascular epidemiology.

When populations are compared, strict attention has to be paid to the methods of sampling. In the present study, people not locally known to be ill were invited to the examination. The definition of "being ill" may vary in different populations and this of course could be a source of bias. It is not believed that

this factor would have caused much difference between East and West Finland, or between the other population groups reported here. Nevertheless, it is felt preferable that in future comparative studies the total population should be sampled, including those manifestly ill.

Geographic Differences.—The preponderance of ECG abnormalities, the difference in the frequency of cases belonging to the "coronary" group as well as in the frequency of hypertensive subjects among the rural group of men from East Finland as compared to the corresponding group from West Finland is in agreement with the results of some other comparisons. In the East, cardiovascular mortality of men is reported to be higher than in the West, pensioned disability due to cardiovascular disease is more frequent, and the serum cholesterol values are higher.[1] It appears that, at least in these populations, the frequency of the above ECG abnormalities and diagnostic items is a population characteristic closely related to the prevalence of cardiovascular disease. Tentatively it may be concluded that those ECG abnormalities are "preclinical" epidemiologic indications of cardiovascular disease. It should be pointed out, however, that this does not imply any statement concerning their prognostic value; the latter must be separately studied.

Definitely abnormal cases certainly occur in the total population more often than in the present series. On the other hand, the mortality of people with "definite" abnormal findings and with manifest illness must be relatively high, and thus the percentage of such subjects in a population may be relatively small. Until studies of total populations are available, no final conclusion can be made of the eventual effects of selection on the frequency of abnormal findings in the present study.

Occupational Differences.—The groups of "directors" and firemen came from the same geographic area, but their occupations differ in many respects. The lumberjacks again represent a rather different occupation, characterized by an exceptionally large energy expenditure.

There was no significant difference in the frequency of cases belonging to the coronary group, though it tended to be higher among "directors" (7.8 per cent) than among firemen (4.8 per cent) and lumberjacks (4.0 per cent). On the other hand, there was a significant difference in the frequency of hypertension between these three groups: 35.9 per cent of the firemen, 25.6 per cent of the "directors," and 11.4 per cent of the lumberjacks had elevated blood pressure ($P < 0.001$). The same difference is seen in the frequency of hypertensive heart disease ($P < 0.01$).

No obvious explanations can be given for these differences. The slight difference in the age distribution between the populations cannot cause the difference. Some of the cases of left ventricular hypertrophy could be explained by the fact that there were several athletes among the firemen. Vigorous training can cause marked increase in the amplitude of the R waves in the chest leads,[9] but this cannot explain the differences in the frequency of hypertension.

The group of lumberjacks may be expected to manifest the effects of prolonged physical activity on coronary heart disease. In agreement with results of some other studies,[10] the frequency of abnormal findings was lowest among the

lumberjacks. The differences between the lumberjacks and the members of the rural group from East Finland were confirmed recently in a study of the total population of the area.[11]

The particularly high frequency of ECG abnormalities among the city firemen, 28.9 per cent, deserves special mention. It was ascribable essentially to hypertensive rather than coronary heart disease. In line with the frequent ECG abnormalities is the observation that the cardiovascular mortality of the Toronto firemen exceeds that of other men in that area.[12] Among possible causes for a high frequency of cardiovascular, and particularly of hypertensive, heart disease among firemen, selection may be of importance. Stocky, muscular men are generally attracted to this occupation, and the same body type is reputed to have a higher than average cardiovascular risk. However, the mode of life and the occupational exposures typical for firemen may perhaps also contribute to the causation of hypertension.

Hypertension and Left Ventricular Hypertrophy.—Hypertension is known to cause left ventricular hypertrophy. Against this background, the finding of high voltage R waves with equal frequency in the different blood pressure groups (Table VI) appears unexpected. Evidently, in an ostensibly healthy working population, the occurrence of high voltage is determined by factors other than "resting blood pressure" as defined in this study. The habitual amount and intensity of exercise may be one of these factors.

SUMMARY

An attempt was made to apply electrocardiography (ECG) as a tool in epidemiologic investigation. Five geographically and occupationally different groups comprising a total of 985 ostensibly healthy Finnish men in the age range from 20 to 59 years were examined. Strictly defined and standardized techniques and criteria were used in the ECG, history, and clinical examination.

Substantial differences occurred in the frequency of ECG abnormalities among the groups with a range from 28.9 per cent among the Helsinki Fire Brigade to 13.1 per cent among lumberjacks of East Finland. The differences were essentially due to left ventricular preponderance, S-T depressions, and T wave abnormalities.

The frequency of coronary involvement varied from 8.9 to 4.0 per cent, with a significant difference between eastern (8.9 per cent) and western (4.1 per cent) rural men. The difference was largely due to angina pectoris and to the category of possible myocardial infarction.

Hypertension also showed significant differences, the range being from 37.2 to 11.4 per cent in the different groups. Correspondingly, hypertensive heart disease varied from 9.3 per cent among firemen to 0 in lumberjacks.

High voltage R waves compatible with left ventricular hypertrophy occurred with the same frequency in men with hypertension, moderately elevated or normal blood pressure.

It is concluded that population differences may be described with the aid of electrocardiography. The causes of the observed differences are discussed.

REFERENCES

1. Keys, A., Karvonen, M. J., and Fidanza, F.: Serum Cholesterol Studies in Finland, Lancet **2**:175, 1958.
2. Roine, P., Pekkarinen, M., Karvonen, M. J., and Kihlberg, J.: Diet and Cardiovascular Disease in Finland, Lancet **2**:173, 1958.
3. Karvonen, M., Orma, E., Keys, A., Fidanza, F., and Brozek, J.: Cigarette Smoking, Serum Cholesterol, Blood Pressure, and Body Fatness. Observations in Finland, Lancet **1**:492, 1959.
4. Karvonen, M. J., Metsälä, P., Pekkarinen, M., and Rautanen, Y.: To be published.
5. Simonson, E.: A Spatial Vector Analyzer for the Conventional Electrocardiogram, Circulation **7**:403, 1953.
6. Blackburn, H., Keys, A., Simonson, E., Rautaharju, P., and Punsar, S.: The Electrocardiogram in Population Studies. A Classification System, Circulation **21**:1160, 1960.
7. Laboratory of Physiological Hygiene, University of Minnesota: Suggested Diagnostic Criteria for Heart Disease in Interpopulation Comparisons.
8. First Report of the Expert Committee on Cardiovascular Disease and Hypertension, WHO Technical Report Series No. 168, 1959.
9. Rautaharju, P. M.: Voltage Changes in the Electrocardiogram as Caused by Vigorous Training, Third World Congress of Cardiology, Abstracts of Communications, Brussels, 1958, p. 411.
10. Keys, A.: Occupational and Environmental Factors in the Development of Heart Disease, *in* Rosenbaum, F. F., and Belknap, E. L., editors: Work and the Heart, New York, 1959, Paul B. Hoeber, Inc., p. 103.
11. Karvonen, M. J., Rautaharju, P. M., Orma, E., Punsar, S., and Takkunen, J.: Heart Disease and Employment. Cardiovascular Studies on Lumberjacks, J. Occup. Med. **3**:49, 1961.
12. Mastromatteo, E.: Mortality in City Firemen. II. A Study of Mortality in Firemen of a City Fire Department, A.M.A. Arch. Ind. Health **20**:227, 1959.

Vol. 19

USE OF COHORT LIFE TABLES IN FAMILY STUDIES OF DISEASE

DAVID J. NEWELL, M.A.*
BERNARD G. GREENBERG, PH.D.
T. FRANKLIN WILLIAMS, M.D., AND
P. BURT VEAZEY, M.D.**
CHAPEL HILL, N. C.

From the Department of Biostatistics, School of Public Health, and the Department of Medicine, School of Medicine, University of North Carolina, Chapel Hill

(Received for publication Oct. 6, 1960)

FAMILY patterns of disease have proved useful in the study of the etiology of certain diseases. They cannot always differentiate between pathways which are genetic, infectious, or due to a common way of life, or which depend on some combination of these three.[1] Nevertheless, they have often provided useful clues, in combination with other information about the disease.

In the search for familial patterns of disease, it is usual to examine relatives of an index case or propositus who has the disease which is under study. This method has proved very fruitful in diseases with an onset early in life. For a disease with a late onset, however, the method will produce a smaller yield of relatives for useful study. The children of the propositus may be too young to show any signs or symptoms of the disease, while the parents of the propositus may be already dead and not available for examination.

It is the purpose of this paper to suggest one way of gathering some information from the parental generation when many of them are dead at the time of diagnosis or study of the disease in their (middle-aged) children. The amount of information about them is often severely limited. The cause of death may be recorded but, as the death may well have occurred at a time when modern diagnostic techniques were not available and registration practices were less reliable than they are today, it would be unwise to put too much faith in the recorded cause. The dates of vital events, however, are usually remembered or recorded, and we limit ourselves to the question "Did the parents of patients with the disease die earlier than would be expected?"

One statistical instrument which can be used in answering this question is the Cohort Life Table.

Supported in part by Public Health Service Research Grant W-74 and Training Grant 2G-38.
*Now at the Medical School, King's College, University of Durham, Newcastle Upon Tyne, England.
**Now at St. Louis City Hospital, St. Louis 4, Mo.

COHORT LIFE TABLES

The underlying idea in the cohort life table is to keep a record of the ages at death of all the people born in a particular year. The mortality rate at each age in the appropriate year is taken as the basis of the table. Thus, for the 1890 cohort, the death rate at age under 1 in 1890 is used, together with that at age 1 in 1891, and so on, until that for age 70 in 1960 is reached. It would be impossible to obtain the actual deaths beyond the present period and, therefore, extrapolations must be used to complete the table as necessary.

As a practical detail, records for the early cohorts are available only every 10 years, and interpolation is used for the intercensal years.

The particular cohort life tables which will be used to illustrate the method in the present paper were prepared in the National Office of Vital Statistics by Sirken.[2] They are based on the Massachusetts Life Table of 1890, and the United States Life Tables for 1900-1902, 1909-1911, 1919-1921, 1929-1931, 1939-1941, and 1949-1951. They give numbers of survivors at 5 and 10 year intervals for three cohorts of 100,000, born in 1890, 1910, and 1930, for white males and females separately. The data for males are shown graphically for age 20 onward in Fig. 1. For females, the curves are similar to those for males, but considerably higher at all ages.

FIRST METHOD—MEDIAN AGE AT DEATH

For each parent of the propositus, we have available the date of birth, the age at which the propositus was born, and the age at death, or present age if still alive. We are not interested in comparing the age at death of the parent with those of all other people born at the same time. The important additional information is that these parents survived the diseases of childhood and lived to generate the propositus. Thus, a man born in 1890 who is the father of the propositus born in 1910 is compared not with the whole 1890 cohort but with that remaining part which had survived until 1910. We shall use the term "cogenerative cohort" to describe people who were born at the same time as the parent of the propositus and were also known to be alive at the time that the propositus was born.

The first way of comparing the age at death of each parent of the propositus with others in the cogenerative cohort is to see whether or not he survived to the *median* age at death of the cohort. The median age at death is based on a French measure of longevity, the "*vie probable*," described in a footnote on page 21 in a paper by Dublin and co-authors.[3] It can most easily be described by an example, using Fig. 1. Of males born in 1890, the proportion who survived to age 20 (the age at the time the propositus was born) can be seen from the graph to be around 70 per cent. The exact survival percentage until age 20 was determined from the numbers used in plotting the graph and was found to be 70.64 per cent. Half of this latter percentage (35.32 per cent) will survive until age 70, and again this can be approximated by referring to the graph. Accordingly, if the fathers of our propositi have no different life span from the rest of their cogenerative cohort, we would expect half of the fathers born in 1890, who were 20 years old when their child was born, to die before the age of 70, and half to survive beyond

70. (The importance of using the cogenerative cohort, rather than the whole cohort, is shown by the fact that half of the whole 1890 cohort had died by age 59, rather than 70 for the cogenerative cohort surviving 20 years.)

To facilitate the computation of the median age at death for each parent, Tables I and II were prepared. The three columns for the 1890, 1910, and 1930 cohorts were first calculated from the cohort survivorship tables. The other columns were then completed by crude interpolation and extrapolation. No very sound basis for extrapolation is available for the years before 1890. The hints contained in the sporadic life tables for Massachusetts from 1850 onward,[3] however, suggest that mortality experience was improving uniformly throughout the whole period and the extrapolations were made on this basis. We assume that the figures for 1850 births also apply to all earlier cohorts.

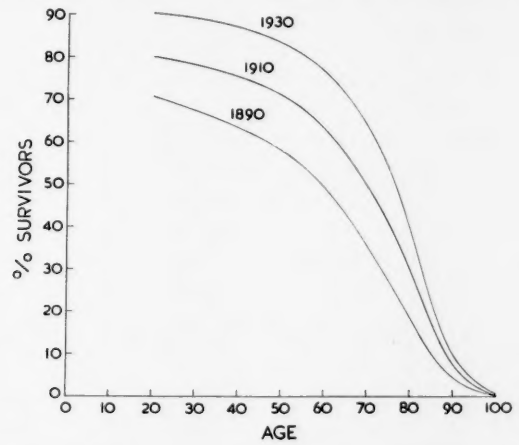

Fig. 1.—Cohort survivorship—Males born 1890, 1910, 1930.

The use of Tables I and II is very simple. The actual age at death (usually available to the last completed year) of each parent of each patient is compared with the entry in the table for the relevant cogenerative cohort and classified as less than, equal to, or greater than the tabular figure. On the null hypothesis that the parents live as long as the general population, equal numbers should live longer or shorter than the median. A binomial test (sign-test) or chi square can be used to determine statistical significance. Results for fathers and mothers should be kept separate; a short life for one sex and a normal length life for the other may give a useful clue to genetic transmission.

Some parents may still be alive at the time of the study and above their relevant median age of death. They can clearly be counted with those who have died above the median age. Parents who are still alive but below the median age present a different problem. It is not legitimate to exclude them, nor is it easy to estimate whether they will survive until the median age. In the example to be given, very few parents fell into this category and there was little difficulty. It there are many such parents in the same cogenerative cohort and there is some idea of the shape of distribution of age at death, one of the methods for

analyzing censored data may be used. These will be discussed at the end of the following section.

TABLE I. MEDIAN AGE AT DEATH OF FATHERS

AGE OF FATHER AT BIRTH OF PROPOSITUS	DATE OF BIRTH OF FATHER								
	1850	1860	1870	1880	1890	1900	1910	1920	1930
20	62	64	66	68	70	72	74	76	78
25	64	65	67	68	71	73	75	76	78
30	65	66	68	69	71	73	75	76	78
35	66	67	69	70	72	74	75	77	78
40	67	68	70	71	72	74	76	77	78
45	68	69	70	71	73	75	76	77	79
50	69	70	71	72	74	76	77	78	79
55	70	71	73	74	75	77	78	79	80
60	73	74	75	76	77	78	79	79	80
65	77	77	78	78	78	79	80	80	81

TABLE II. MEDIAN AGE AT DEATH OF MOTHERS

AGE OF MOTHER AT BIRTH OF PROPOSITUS	DATE OF BIRTH OF MOTHER								
	1850	1860	1870	1880	1890	1900	1910	1920	1930
20	71	72	73	74	76	78	80	81	82
25	73	74	75	76	77	79	81	81	82
30	73	74	75	76	77	79	81	81	82
35	73	74	75	76	77	79	81	81	82
40	74	75	76	77	78	80	81	81	82
45	74	75	76	77	78	80	81	81	82
50	75	76	77	78	79	81	82	82	82

SECOND METHOD—EXPECTATION OF LIFE

The foregoing method has compared each parent's age at death with an average figure, without taking account of the magnitude of the difference between the two. In order to make use of the additional information, it is necessary to use a different average figure—the mean age at death rather than the median age at death.

Once again, each parent is referred to his cogenerative cohort. For each cohort, the usual life table calculations give the expectation of future life at each age. The distribution of years of future life is illustrated for certain ages in Fig. 2.

By adding this expectation of future life at the time of birth of the propositus to the age then achieved, we find the expected mean age at death of the parent. Thus, for men of age 20, the distribution of future life is negatively skewed or may be thought of as having a long tail on the left-hand side of the distribution of future life. At 40 the skewness is less marked, by 60 the skewness has become positive, and by 80 the distribution is J shaped or exponential.

A table similar to Table I was prepared, giving mean ages at death to replace the median ages. The method involved calculation of expectation of life for the 1890 cohort (Table III) and the 1910 cohort. Mean ages for earlier cohorts were obtained by linear extrapolation. The result is shown in Table IV. The two tables were almost identical for ages 50-60; but, for fathers who were young at the time of birth of the propositus, the mean was smaller than the median, being as much as 5 years earlier in the 1850 cohort.

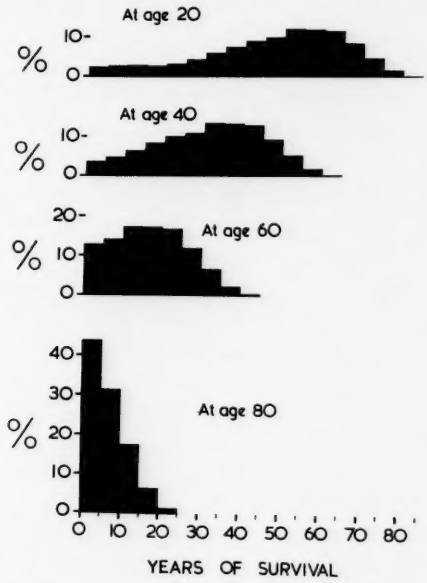

Fig. 2.—Percentage distribution of years of remaining life—U. S. white males, 1890 cohort.

If all of the parents are dead, it is easy to calculate for each parent the difference between age at death and the mean age at death of the cogenerative cohort. On the null hypothesis, the population mean of these differences is zero, and the mean of the observed differences can be compared with zero by a t test, which is "robust" to a certain amount of non-normality.

The problem of parents who are alive at the time of the study is more acute in the comparison with mean age than in the simpler comparison with median age. In the section "example" hereafter it will be seen that most of the living parents had already exceeded their median age at death and could thus be classified in the first method. In the second method, on the other hand, the magnitude of the difference between the observed age at death and that expected is unknown for *all* live parents.

The solution to this problem probably lies in the developing field of study of censored samples. At the present stage, estimates are available for the means of censored data from normal distributions[4-6] and from exponential distributions.[7] Unfortunately, with the possible exception of ages 40 and 80, few of the cohorts' distributions of ages at death are fitted by such distributions (Fig. 2) and the robustness of these methods to non-normality is not yet known.

TABLE III. MEAN AND MEDIAN AGES AT DEATH OF THE 1890 MALE COHORT

AGE X	$\overset{o}{E}_x$	MEAN AGE AT DEATH	MEDIAN AGE AT DEATH
20	46.8	66.8	70.1
25	42.9	67.9	70.6
30	39.0	69.0	71.2
35	35.2	70.2	71.7
40	31.1	71.1	72.3
45	27.2	72.2	73.1
50	23.4	73.4	73.8
55	20.0	75.0	75.0
60	16.8	76.8	76.6
65	14.0	79.0	78.4
70	11.2	81.2	80.5
75	8.9	83.9	83.1

The "expectation" figures are derived from the formula based on Simpson's rule:

$$T_x = \frac{5}{3}(l_x + 4l_{x+5} + 2l_{x+10} + 4l_{x+15} + 2l_{x+20} + \ldots.)$$

used in the Abridged Life Tables for Wales and London,[12]

and $\overset{o}{e}_x = T_x/l_x$; l_x is plotted in Fig. 1.

TABLE IV. MEAN AGE AT DEATH OF FATHERS

AGE OF FATHER AT BIRTH OF PROPOSITUS	DATE OF BIRTH OF FATHER						
	1850	1860	1870	1880	1890	1900	1910
20	57	59	62	64	67	69	72
25	59	61	64	66	68	70	72
30	61	63	65	67	69	71	73
35	63	65	67	68	70	72	74
40	65	67	68	70	71	73	74
45	67	68	70	71	72	74	75
50	68	70	71	72	73	75	76
55	70	72	73	74	75	76	77
60	73	74	75	76	77	78	79
65	76	77	78	78	79	80	81

SHORTCOMINGS OF THE TECHNIQUES

Before considering the application of the methods outlined above, it would be well to look at their shortcomings.

1. Cohort life tables, as constructed at present, are not what they purport to be in that they assume that the yearly total death rates in the area studied are relevant to the original cohort of people born. This ignores the fact that new migrants to the area are different people and may have different mortality rates. For example, Gordon[8] shows that Japanese immigrants to the United States have favorable mortality rates when compared with the total white population, and Dublin[9] has shown that, in general, the foreign born have a lower expectation of life than the native born. While this is a criticism of the tables in general,

it may be a fault in the right direction for this application, since the parents of the propositi include migrants.

2. Year-to-year fluctuations in intercensal periods are ignored in the present tables. These may be considerable in times of major wars or epidemics. Dublin, Lotka, and Speigelman[3] state on page 176 of their book that any attempt to correct for such fluctuations would involve an amount of labor out of proportion to the effect on the conclusions to be drawn.

3. In the construction of the tables, mortality rates are extrapolated to future years. This should not be of much importance to us, since we are mainly concerned with cohorts whose lives are completed. In the present study, however, the lack of data for earlier years has made it necessary to extrapolate backward from figures which have been constructed by forward extrapolation. In the future, this problem will cease to exist, when the parental generation has adequate records of vital statistics.

4. The method outlined above takes account of the fact that the parents have survived until the age at which the propositus was born. It does not, however, take account of the fact that the parents have actually produced a child. It is well known that *marriage* is associated with favorable mortality experience and, presumably, although data are not so easily adduced, *parenthood* also carries with it the chance of a longer life than does infertility (in the demographic sense of the word). In view of the changes in marital status of an individual throughout life, it seems unlikely that cohort life tables for those who have borne or fathered a child will become available, unless a much more detailed form of vital registration is introduced (see Dublin, Lotka, and Speigelman[3] [page 135] for an account of this difficulty).

5. Finally, the cohort life table used refers to the populations of Massachusetts and of the whole of the United States at different periods, whereas the patients in a particular study will generally come from a much more restricted area, with possible more or less favorable mortality experience.

It is not at present possible to say to what extent the above shortcomings affect the numerical values which should be used for median and mean expectation of life. Until the method is more fully understood, we recommend the use of one or more control samples in each study.

EXAMPLE

A randomly selected group of 248 diabetic patients aged over 35 years, who had attended the Memorial Hospital, Chapel Hill, North Carolina, during a period of 30 months, was chosen for study. Mailed questionnaires were sent to each propositus, asking for the relevant dates of birth and death of parents and propositus. The total number of completed returns after a follow-up letter was 162, a yield of 65.3 per cent of those questioned.

Owing to the possible bias caused by questionnaires unreturned at the rate of one in three, the representativeness of the propositi completing the questionnaire was investigated. Among males, the average ages of respondents and non-respondents were identical in both the private and staff patients. Moreover, the ages of the immediate respondents and those replying after the follow-up letter

showed no differences whatsoever. (This similarity between original respondents and follow-up letter respondents was repeatedly observed in all other characteristics besides age which were investigated and, hence, no further distinction will be made between the two groups of respondents.) Among females, the respondents were 3 to 5 years younger than the nonrespondents but this was not of sufficient magnitude to be statistically significant.

The response rates themselves showed no difference between the private and staff patients. The females tended to reply at a slightly greater rate than males, 71.7 per cent versus 60.9 per cent. This difference in the response rates of the sexes was not statistically significant; moreover, its role in causing any bias would be only relatively minor, if important at all.

Based upon the nonsignificant differences in age, sex, pay status, and whether the propositus replied to the original questionnaire or its follow-up letter, the inference was made that the one-third nonrespondents would not introduce a bias in studying the available returns. Of the 162 replies available, only 151 were able to be used because they provided the necessary data. In a prospective study, a higher yield could be expected.

The first method of comparing the fathers' age at death with the median age is summarized in Table V. Of the fathers already dead, 64 died before the median age and 63 after.* Of those still alive at the time of the study, 13 had already passed the median age, and 4 had not reached it. We have no simple rule for allocating these last 4 to one group or the other. All 4 of them had at least 10 years to live before achieving their median age, and there were not sufficient fellow members of their cohort to make a good estimate of their future life, with the use of censored data techniques.

TABLE V. COMPARISON OF THE AGE AT DEATH OF FATHERS OF DIABETICS WITH THE MEDIAN AGE FOR THE COGENERATIVE COHORT*

DATE OF BIRTH OF FATHERS	FATHERS WHO DIED		FATHERS ALIVE	
	BEFORE MEDIAN	AFTER MEDIAN	BEFORE MEDIAN	AFTER MEDIAN
1828–1849	1	7	0	0
1850–1859	11	15	0	0
1860–1869	21	22	0	0
1870–1879	20	15	0	7
1880–1889	7	4	0	6
1890–1901	4	0	4	0
Total	64	63	4	13

*In addition, 7 fathers died in the same year as the median age for their cogenerative cohort.

At least 64 of the fathers will have died before the median and, at most, 80 will have survived beyond it. That is to say 64/144 = 44.4 per cent is the best

*There were 14 cases where either the mother or father of the propositus was less than 20 years of age at the time of birth of the propositus. Since Tables I to IV start at age 20, the mean or median associated with the latter figure was used for those parents only 1 or 2 years younger than 20. For those parents younger than 18, a crude extrapolation was made in the appropriate table.

estimate of the proportion below the median, if all 4 remaining survive beyond the median. The standard error of this proportion is 4.1 per cent, and 95 per cent confidence limits are 36.3 per cent to 52.5 per cent.

Alternatively, if we suppose that all 4 of the surviving fathers are to die before the median age, the comparison would be between 68 early deaths and 76 late deaths. This gives a proportion of 47.2 per cent with a standard error of 4.2 per cent and confidence limits 39.0 per cent to 55.4 per cent.

Whether one uses the 64 or 68 early deaths, the resultant confidence band for the percentage below the median includes the value of 50 per cent, a value hypothesized for the null case. Since neither case would be significant, we can retain the null hypothesis. We have no reason for asserting that the fathers of the propositi have a different length of life from that of their cogenerative cohorts. In fact, we were expecting to demonstrate a shorter life span for the fathers of the propositi, and a one-sided test of significance would have demonstrated more convincingly that we should not reject the null hypothesis.

In Table V, the results are subdivided according to the date of birth of the father. This provides a crude check on the validity of the cohort life tables used. Two points can be made from this table. The first is that in the earliest cohorts there is a (statistically nonsignificant) tendency for fathers to live longer than the median. The second point is that in the most recent cohort of fathers, the only deaths that can be recorded are the early ones, since the median point had not yet been reached. Nevertheless, for the 17 fathers still alive, the information necessary for this method of analysis is absent in only 4 cases. If the analysis is restricted to fathers born before 1890, the analysis is much simpler, at the expense of losing information on all 8 of the fathers born after that date.

When we turn to the second method of analysis, however, information is missing for all 17 of these men. Each father's age at death or survival was compared with the corresponding value in Table IV, with the results shown in Table VI. We do not know the best way of analyzing Table VI. Unless the survivors all live to a great age, the final distribution, when all the survivors have died, will have a marked negative skewness (as would be expected from Fig. 2, and what is known about the age at fatherhood). If they had all died immediately after the questionnaire was returned, the average difference between recorded age at death and the expected mean age at death for the 17 fathers would have been +8.24 years. For the complete sample, the corresponding figure would have been +1.69 years, which is not significantly different from zero. Since we have never entertained the possibility that the parents of diabetics live *longer* than expected, a one-tailed test of significance is relevant. If some of the survivors continue to survive, the mean difference of 1.69 years would increase, and, on a one-tailed test, significance would have less chance of being established.

If the numbers in the study were larger, we could use an actuarial method to analyze Table VI. This would assume that the survivors would die at the same rate as members of previous cohorts of comparable age, after adjustment for year of birth and age at birth of child. Such methods and significance tests for them have been discussed by Chiang.[10,11]

TABLE VI. COMPARISON OF AGE AT DEATH OF FATHERS OF DIABETICS WITH THE MEAN AGE FOR THE COGENERATIVE COHORT

OBSERVED AGE MINUS EXPECTED AGE AT DEATH (YEARS)	FATHERS ALREADY DEAD	FATHERS STILL ALIVE
−37 to −33	2	
−32 to −28	3	
−27 to −23	7	
−22 to −18	3	
−17 to −13	6	
−12 to −8	9	1
−7 to −3	18	3
−2 to 2	19	0
3 to 7	20	3
8 to 12	21	3
13 to 17	16	4
18 to 22	5	3
23 to 27	4	
28 to 32	1	
Total	134	17
Mean	+0.86 years	+8.24 years

THE CONTROL SAMPLE

A control sample double the size of the diabetic sample was randomly chosen from ex-hospital inpatients without diabetes. They were matched with the diabetics by age (in decades), sex, and pay status at the hospital. All members of both samples were white. A questionnaire was mailed to all controls and 52 per cent replied. A follow-up letter increased the yield to 73 per cent of usable replies. Again, the question of bias was investigated to determine whether the control group could be used as a random segment of the population sampled as well as to answer the question whether it was still a control for the diabetic propositi.

The distributions of age, sex, pay status, and rate of replying to the questionnaire were compared between the respondents and nonrespondents. As was observed among the diabetic propositi, there were no significant differences between those who replied and those who failed to do so. The slight tendency of more of the females to reply to the questionnaire was noted again but the difference this time was smaller, 75 per cent versus 70 per cent. Although the sample size among the controls was twice that of the diabetics, the difference in response rates between the two sexes was again not of sufficient magnitude to be considered other than a chance happening. Hence, the first question could be answered by reporting that the completed questionnaires appeared to be a random selection from the population sampled.

The control questionnaires which were returned had age, sex, and pay status distributions identical with those observed among the questionnaires completed by the diabetic propositi. This assured the investigators that a comparison of the results from the control questionnaires could be made validly against those of the diabetic group.

The control sample can be used in two ways. The first and obvious way is to make a direct comparison of the diabetic sample with them. The second way, which will be described next, is to test whether the cohort generation life tables can be used for patients of this hospital. The method of analysis corrects for age and sex, and the remaining question is whether the table can be applied to people of the same pay status as the diabetics, from the geographic area served by the hospital.

Of the replies, 364 gave all the necessary information about the father of the control. In Table VII, the age at death is compared with the median.* Thirty-three of the fathers were alive at the time of the questionnaire, 12 of whom had not reached the median age. If these 12 die before the median, the comparison will be between 154 early deaths and 195 late deaths, this difference from equality would arise by chance less than once in 20 times (chi square = 4.58, 1 degree of freedom). If all the present survivors live beyond the median, the comparison would be between 142 and 208; this difference is highly significant, and would arise by chance less than once in a thousand times (chi square = 12.1, 1 degree of freedom).

TABLE VII. COMPARISON OF AGE AT DEATH OF CONTROL FATHERS WITH THE MEDIAN AGE FOR THE COGENERATIVE COHORT*

DATE OF BIRTH OF FATHERS	FATHERS WHO DIED		FATHERS ALIVE	
	BEFORE MEDIAN	AFTER MEDIAN	BEFORE MEDIAN	AFTER MEDIAN
1824–1849	9	18		
1850–1859	19	40		
1860–1869	38	73		
1870–1879	41	35		8
1880–1889	28	8		12
1890–1901	7	1	12	
Total	142	175	12	20

*In addition, 14 fathers died in the same year as the median age for their cogenerative cohort, and 1 was still alive at exactly the median age.

If the comparison is made for fathers born before 1890, 135 of these died before the median age and 194 after, none of these fathers being unclassifiable into one of the two groups. This difference from equality is significant at the 1 per cent level (chi square = 10.2, 1 degree of freedom).

Table VIII shows the result of comparing the actual age at death of the control fathers with the mean as given by Table IV. On average, fathers already dead had lived 2.30 years longer than would be expected from the life table, and the 33 still alive had at the time of the survey exceeded their expectancy by 5.15

*There were 29 cases where one of the parents of the control propositus was less than 20 years of age at the time of birth of the propositus. The procedure outlined in a previous footnote for treating this occurrence was followed in the control group. The rate of occurrence of young parents among the control propositi was identical to that of the diabetic group, and this provided further assurance about comparability of the samples.

years. If we assume that the 33 all died immediately, the whole sample would have lived 2.55 years longer than expected. A fortiori, if the survivors live for some time after the survey, the whole group will live for a longer average time than expected.

TABLE VIII. COMPARISON OF AGE AT DEATH OF CONTROL FATHERS WITH THE MEAN AGE FOR THE COGENERATIVE COHORT

OBSERVED AGE MINUS EXPECTED AGE AT DEATH (YEARS)	FATHERS ALREADY DEAD	FATHERS STILL ALIVE
−42 to −38	4	
−37 to −33	5	
−32 to −28	1	
−27 to −23	5	
−22 to −18	10	
−17 to −13	20	
−12 to −8	17	6
−7 to −3	46	2
−2 to 2	44	6
3 to 7	49	7
8 to 12	56	2
13 to 17	44	5
18 to 22	15	3
23 to 27	13	2
28 to 32	2	0
Total	331	33
Mean	+2.30 years	+5.15 years

For this control group, then, both the median ages in Table I and the mean ages in Table IV are too pessimistic. There are several possible reasons for this discrepancy. First, the extrapolations were based on the assumption of linearity of change of median or mean age at death with increasing birth year. In this group, the rate of change before 1870 appears to be less than that in the years 1890-1930. Second, the white population of this area of the country may have a longer life than the populations on which the tables were based. Third, diabetes tends to attack the higher socioeconomic classes, who have a longer life expectancy.* Since our control sample was chosen to match the sample of diabetic individuals in socioeconomic level, this factor would have affected both groups. Fourth, extravagant claims of longevity may have been made by people born before vital registration documents were obligatory.† Finally, the response to the mailed questionnaire may have been biased toward the more literate members of the sample, who may well have a higher life expectancy in their families.

Since the cohort life tables by themselves were not quite appropriate, the study in this example was completed by a comparison between the fathers of the

*Diabetes is associated with obesity which, in turn, is related to social class and a way of life. Moreover, higher socioeconomic classes have a longer life expectancy which enables them to have greater exposure to diabetes with late onset. That is, the crude rate of diabetes may be related to social class because of variations in life expectancy even though the net probability of risk may not be so related.

†One of our patients, alive in 1960, claimed that his father was born in 1828, the patient himself being born in 1887 when his father was 59. This is by no means impossible, but we are interested in recording a span of at least 132 years by two generations.

patients and the controls. This comparison used the cohort life tables as a method of controlling for date of birth. By design, the two groups were matched for date of birth of the propositus. Analysis showed the distributions of dates of birth of fathers not to differ significantly between the two groups. Thus the groups were comparable in so far as we were able to judge. Further analysis showed the proportions of fathers still alive at the time of the study to be 11 per cent for diabetics and 9 per cent for controls. These figures do not differ significantly. Finally, analyses were made of the duration of life in those already dead and in the survivors, separately. For those already dead, the fathers of diabetics had lived 0.86 years longer than expected, and the fathers of controls 2.30 years longer.

The standard error of this difference (of 1.44 years) is 1.38 years, showing no significant difference. For the survivors, the fathers of diabetics had already lived 8.24 years longer than expected. These figures also give no evidence to support the hypothesis of a shorter life span for the fathers of diabetics.

The conclusion from the example is that the fathers of diabetics have no evidence of a different life span from that of matched control patients with other diseases. Similar results were obtained for the mothers.

SUMMARY

1. The age at death of parents can be a useful clue in the study of the etiology of diseases. Until recently, cohort life tables suitable for studying normal values have not been available.

2. The cohort life table can be used to provide average values (mean and median) for the length of life of parents born at a particular time and known to have survived until the birth of the propositus under study.

3. The shortcomings of the method are discussed in this paper and refinements which will become possible in the future are pointed out.

4. The method is applied to a study of diabetes in middle-aged white persons. It is shown that there is no evidence to suggest that their parents have shorter lives than the general population or than a control group matched for age, sex, and socioeconomic status.

The cohort life tables for cohorts born in 1890, 1910, and 1930 were compiled by Dr. Monroe Sirken of the National Office of Vital Statistics, Washington, D. C. Permission for their use in this paper was granted by Dr. Halbert Dunn, Chief of that Office.

REFERENCES

1. Lilienfeld, A. M.: A Methodological Problem in Testing a Recessive Genetic Hypothesis in Human Disease, Am. J. Pub. Health **49**:199, 1959.
2. Sirken, M.: Cohort Generation Study, National Office of Vital Statistics. (Unpublished communication, 1959.)
3. Dublin, L. I., Lotka, A. J., and Speigelman, M.: Length of Life, New York, 1949, Ronald Press Co.
4. Sarhan, A. E., and Greenberg, B. G.: Estimation of Location and Scale Parameters by Order Statistics From Singly and Doubly Censored Samples. Part I, Ann. Math. Stat. **27**:427, 1956.
5. Sarhan, A. E., and Greenberg, B. G.: Estimation of Location and Scale Parameters by Order Statistics From Singly and Doubly Censored Samples. Part II, Ann. Math. Stat. **29**:79, 1958.

6. Cohen, A. C., Jr.: Simplified Estimators for the Normal Distribution When Samples Are Singly Censored or Truncated, Technometrics **1**:217, 1959.
7. Sarhan, A. E., and Greenberg, B. G.: Tables for Best Linear Estimates by Order Statistics of the Parameters of Single Exponential Distributions From Singly and Doubly Censored Samples, J. Am. Stat. A. **52**:58, 1957.
8. Gordon, T.: Mortality Experience Among the Japanese in the United States, Hawaii and Japan, Pub. Health Rep. **72**:543, 1957.
9. Dublin, L. I.: The Longevity of the Native Born and the Foreign Born in the Population of the United States. International Congress of Population at Paris, 1937, Paris, 1938, Mermann & Cie, vol. 5, pp. 188-194.
10. Chiang, C. L.: A Stochastic Study of the Life Table and Its Applications. II. Sample Variance of the Observed Expectation of Life and Other Biometric Functions, Human Biol. **32**:221, 1960.
11. Chiang, C. L.: A Stochastic Study of the Life Table and Its Applications. III. The Follow-Up Study With the Consideration of Completing Risks, Biometrics. (To be published.)
12. Registrar General of England and Wales: Life Tables, Decennial Supplement 1951, London, 1957, Her Majesty's Stationery Office, p. 37.

Vol. 1
196

MULTIDISCIPLINARY STUDIES OF ILLNESS IN AGED PERSONS

V. A New Classification of Socioeconomic Functioning of the Aged

THE STAFF OF THE BENJAMIN ROSE HOSPITAL*
CLEVELAND, OHIO

From the Benjamin Rose Hospital and from the Departments of Medicine and Preventive Medicine, School of Medicine, Western Reserve University and the University Hospitals

(Received for publication Aug. 20, 1960)

ONE of the aspects of aging is the process of mental, physical, and social deterioration of an individual. Specifically defined, quantitative descriptions of the deterioration are lacking. Such information is, however, urgently needed if we are to evaluate means of delaying the aging process or modifying it for the benefit of the individual and society. In recent years, the staff of the Benjamin Rose Hospital, a hospital for prolonged illness in Cleveland, Ohio, has been studying mental, physical, and social changes in selected aging persons in order to evolve suitable definitions and methods for obtaining the indicated quantitative descriptions of deterioration.[1,2,3] An over-all *Index of Independence in Activities of Daily Living* (*Index of ADL*) was presented earlier as a more specific index than previously available, and its usefulness was demonstrated.[2,3]

In preliminary searches for indices of social deterioration, studies were made of changes in nature of residence and changes in nature of economic support.[1,3] Thus, for example, a change in residence from a single home, double home, apartment, or hotel to a nursing home, hospital, or institution of custodial type was considered to indicate deterioration. A change in economic support from a

This investigation was supported by Research Grant 7200 made to Western Reserve University School of Medicine by the United States Public Health Service for an interdisciplinary program in research on aging.

*Sidney Katz, M.D. and Austin B. Chinn, M.D. (principal investigators); Arthur S. Littell, Sc.D. (biostatistician); Amasa B. Ford, M.D., Kingsbury G. Heiple, M.D., Malcolm S. Mackenzie, M.D., William B. Newberry, M.D., and Alexander P. Orfirer, M.D. (medical, surgical, and psychiatric investigators); Marjorie W. Jaffe, B.A., M.A. (co-ordinator); Ruth E. Mason, A.B., M.S.S.A., and Ruth M. Jenkins, A.B. (social service investigators); Marion Bittman, B.A., M.N., Georgina Greene, M.S., R.N., Lavina M. Robins, B.S., R.N., H. Jean Holman, R.N., Mary A. Sherback, R.N., and Jacquelyn Millen, R.N. (nursing investigators); Susan W. Ritchie, O.T.R., and Stephen J. Takacs, R.P.T. (occupational and physical therapy investigators).

condition of self-support to one of family or community support was similarly defined as indicating social deterioration. Although such definitions were useful in defining certain instances of social deterioration, preliminary studies gave us the impression that there were more instances, but they were obscured by certain modifying factors not taken into account in these definitions. We studied the case histories of study patients in search of such factors and found that they did exist. Certain patients, for example, required the types of assistance given in nursing homes, but by virtue of their plentiful economic resources could avoid going into nursing homes by employing enough assistance to establish the equivalent of a nursing home situation within their own residences. It thus appeared that social deterioration should be defined in relation to individual *resources* and *needs* as well as in relation to the more restricted observations of changes in residence and nature of economic support. It was also obvious that change in employment status or *productivity* was a characteristic of socioeconomic deterioration in certain persons.

The present report defines an over-all *Index of Independence in Socioeconomic Functioning* (*Index of ISF*), specifically defined and newly evolved from a multivariate analysis of detailed information concerning patient resources, needs, and productivity. The nature of the information and the analysis are presented in the Methods section that follows. Illustrations are given of its uses in describing changes relative to chronic disease and chronic disability of aged persons. The economic factors in the index were derived in our own independent analysis, but their inclusion is consistent with the nature of important economic factors previously reported by Steiner[4,5] and considered in the Discussions section of this paper.

METHODS

The *Index of Independence in Socioeconomic Functioning* (Table I) was evolved from data obtained in initial and repeated follow-up examinations made at 6 month intervals during the course of illness. In the earlier studies, information was obtained on all patients admitted with fracture of the hip which occurred within 6 months prior to admission. The study was extended subsequently to include all patients admitted with a first cerebral infarction occurring within 30 days prior to admission. The methods of accumulating and recording physical, psychological, and social data were described previously.[1] We wish to re-emphasize that, as part of the rehabilitation aim in The Benjamin Rose Hospital, patients are urged to function at their most independent levels. They are, in fact, taught to function more independently when possible and are tested at more independent levels than they demonstrate. It is recognized that in certain instances the level of a given patient's functioning may not truly reflect his ability as, for example, in the event of poor motivation or outright refusal; however, tests that eliminate these limiting factors are not available currently. By making observations in this setting where efforts are directed toward achieving maximum independent functioning and by making direct observations in the home as well as by questions, we attempted to get accurate measures of functional characteristics. At the same time, functional characterization of patients on the basis of

TABLE I. INDEX OF INDEPENDENCE IN SOCIOECONOMIC FUNCTIONING*

	AGENCY SUPPORT	ASSISTIVE-TYPE RESIDENCE		RESIDENCE OWNERSHIP	EMPLOYMENT STATUS
Class I	+	+		+	+
Class II	+ +	+ +	or	+ 0	0 +
Class III	+ +	+ 0	or	0 +	0 0
Class IV	+ 0	0 +	or	0 0	0 0
Class V	0	0		0	0
Class other	Any other pattern of 0's and +'s				

*"+" means that the person is in the "better" category of this item, and "0" means that the person is in the "poorer" category of this item. See text under Methods for detailed definitions of functional status in relation to individual items.

status rather than on the basis of impressions of ability minimized the effects of subjective interpretation. Thus, for example, a person observed to function at a given level was classed as functioning at that level even though there might be some feeling that he was capable of some higher level of functioning.

In order to develop an ordered scale of socioeconomic functioning, information concerning patient *resources*, *needs*, and *productivity* was obtained by direct observation and by a series of specific questions asked of every patient or his family. Information was obtained in the hospital and on routine follow-up visits to the home. The following items were recorded for each patient: (1) type of residence (single home, double home, apartment, hotel, boarding house, custodial residence, nursing home, or hospital); (2) ownership of residence (patient or spouse legal owner or not); (3) gainful employment (number of hours employed in the last 2 weeks); (4) personal living assistance in relation to bathing, dressing, going to toilet, feeding, preparing meals, cleaning house, and washing clothes as previously defined[2]; (5) persons giving daily living assistance (specifically designated family members, attendants, companions, housekeepers, institutional staff, friends, visiting nurse, or other); (6) sources of income (specifically designated family members, agencies of charitable type, private industrial plans, insurance, retirement plans, savings, salary, stocks, bonds, or Social Security); and (7) goods and services purchased (food, clothing, residence tax or rent, and medicines and doctors' fees).

The next step was to subdivide each item into two grades, "better" and "poorer," according to whether the entry described a relatively independent or a relatively dependent level of socioeconomic functioning. Thus, for example, a patient who lived in a house owned by himself or his spouse was defined as one functioning in a "better" category than a patient living in a house which he or

his spouse did not own. Upon completing the grading of all items, an ordered sequence of the graded items or groups of items was sought such that individuals profiled according to this ordered sequence could be ranked relative to each other. An ideally ordered, hypothetical sequence of items or groups of items would permit comparable ranking of all patients; thus, all patients would fit into one of the following profiles:

	ITEM A	ITEM B	ITEM C	ITEM D	
Profile I	+	+	+	+	
Profile II	+	+	+	0	
					etc.
Profile III	+	+	0	0	
Profile IV	+	0	0	0	
		etc.			

where "+" means that the person is in the "better" category of this item and "0" means that the person is in the "poorer" category of this item.

As might be expected, the initial attempt to group and grade specific individual items did not result in an ordered sequence of profiles by which all patients could be ranked relative to each other, such as the pattern of profiles postulated above. By modifying the grouping of items and by changing the dividing line between grades of status for individual items, it was possible to increase the number of patients fitting into the desired pattern of profiles. As this was being done, attention was given to the additional aim of having the index longitudinally sensitive to change from one time to another. This purpose was accomplished by a similar method of weighting definitions of items, groups of items, and grades after having tabulated all changes occurring from one time to another in relation to each item.

The index that was finally evolved appeared to be the best possible on the basis of the available information and will be defined in the section on Results. Items used in the current index narrowed down in number to include status in agency support, assistive-type residence, residence ownership, and gainful employment.

Patients were classed in a "poorer" category relative to agency support if they received financial assistance from agencies of charitable type toward any of the following specific goods and services: food, clothing, residence tax or rent, medicines, and doctors' fees. Aid for the Aged, private social agencies, and government agencies were included among the types of charitable agencies, while Social Security, family assistance, and retirement plans such as industrial and teachers' retirement plans were not included in this category. Patients were classed in a "better" category relative to agency support if they received none of the previously indicated financial assistance from the designated types of charitable agencies.

Two grades of assistive-type residence were evolved for the final form of the index, the "better" category being termed a nonassistive or minor assistive-type residence and the "poorer" category being termed a major assistive type residence. Patients were classed as living in major assistive-type residences if they were living in any type of residence in which they received supervision or active personal assistance in all the 3 functions of bathing, dressing, and going to toilet as defined previously, or if they were living in hospitals, custodial institutions, or nursing-type homes.[2] Patients not receiving supervision or personal assistance in all 3 functions and not living in hospitals, custodial institutions, or nursing-type homes were classed as living in nonassistive or minor assistive-type residences.

Patients who were the legally designated owners of the residences in which they lived were classed in a "better" category relative to this item than patients who were not legally so designated. In this item, the nature of the functional residence was the basis for grading; thus, a patient living in a residence that he did not own was classed in a "poorer" category even though he may have been the legally designated owner of some other dwelling, while a patient living in a residence that he did own was classed in a "better" category even though someone else may have been paying his property taxes.

Gainfully employed patients were graded as socioeconomically more independent relative to this item than patients who were not gainfully employed; thus, the former were considered to be in a "better" category, while the latter were classed in a "poorer" category. Patients were considered as gainfully employed if they were paid for an activity performed regularly at least 1 hour per week, including even such activities as baby sitting. Receiving income such as rent, dividends, or interest was not in itself considered as gainful employment.

RESULTS

A total of 359 evaluations made at 6 month intervals in 93 patients admitted with fracture of the hip served as the basis for evolving the *Index of Independence in Socioeconomic Functioning*. During the course of these evaluations, 21 of the patients died. The 359 evaluations were made during life and included 93 made prior to illness by history, 87 made 6 months after fracture, 78 made 1 year after fracture, 56 made 18 months after fracture, 33 made 2 years after fracture, and 12 made 30 months after fracture. The sample included 6 men and 87 women, ranging in age from 50 to 98 years at the time of fracture. A description of the sample according to age, sex, and marital status is presented in Table II. It is noteworthy that the proportion of widows increased, and the proportion of women with living spouses decreased, with increasing age.

In the multivariate analysis described under Methods and used to evolve the *Index of Independence in Socioeconomic Functioning*, the following patterns were noted in 324 of the 359 evaluations. When socioeconomic assistance as previously defined was received in relation to one function alone, that assistance was in relation to the patient's employment status. When the assistance was in relation to 2 functions, the functions were employment and residence ownership.

The occurrence of 3 spheres of assistance was in relation to employment, residence ownership, and assistive-type residence, while the occurrence of 4 spheres of assistance was in relation to the previous 3 plus agency support. In 24 of the 25 exceptions to this pattern, the observed pattern differed in only one function, and that function was the next function in the ordered sequence. On the basis of this analysis, the following scale was devised by means of which patients could be ranked relative to each other and from one time to another (Table I). The ordered scale was termed the *Index of Independence in Socioeconomic Functioning* (*Index of ISF*) and included 5 classes, individual class items being defined in the Methods section. Class I included persons independent of agency support, living in nonassistive or minor assistive-type residences, owners of the residences in which they lived, and employed gainfully. Class II included persons who differed from those in Class I in only one respect, namely, that they were either not employed gainfully or did not own their own places of residence (not both). Persons in Class III were not employed gainfully and were either not the owners of their places of residence or lived in major assistive-type residences (not both). Persons in Class IV were not employed gainfully, did not own their places of residence, and either lived in major assistive-type residences or received agency support (not both). Persons in Class V were not employed gainfully, did not own their places of residence, lived in major assistive-type residences, and received agency support.

Application of the *Index of Independence in Socioeconomic Functioning* to a few specific examples will serve to leave the reader with a more graphic impression of its meaning. The following illustrations are, therefore, presented.

TABLE II. AGE, SEX, AND MARITAL STATUS OF SAMPLE OF AGED PATIENTS WITH FRACTURE OF THE HIP IN WHOM THE *Index of Independence in Socioeconomic Functioning* WAS EVOLVED

	AGE RANGE (YEARS)					TOTAL
	50–59	60–69	70–79	80–89	90–99	
Women						
With living husband	3	3	3	4	0	13
Widowed	0	4	15	24	8	51
Never married, separated, or divorced	3	3	9	8	0	23
Men						
With living wife	1	0	1	2	0	4
Widower	0	0	0	1	0	1
Never married	0	1	0	0	0	1

Dr. D., a 90-year-old retired chiropractor, lived in an apartment that he owned, was functionally independent in activities of daily living, and received no agency or community financial support. His level of socioeconomic functioning was thus classed as II by the *Index of ISF*. Six months after experiencing a "stroke," he was back in the apartment that he owned, still not gainfully employed and still financially independent. He was, however, now being assisted in bathing, dressing, and going to toilet by a practical nurse employed by his wife. His level of socioeconomic functioning was accordingly Class III by the *Index of ISF*. A year after the onset of his illness,

the practical nurse was no longer in the setting, and Dr. D. was bathing, dressing, and toileting himself in the same economic and living situation; thus, he was again functioning at a Class II level.

Mr. C., a 64-year-old executive, lived in a single home that he owned, was financially independent, and was functionally independent in his activities of daily living. His level of socioeconomic functioning was, thus, I by the *Index of ISF*. Six months after an illness, he was functioning independently in activities of daily living. Although he was financially independent as before, he had not yet returned to work, and his level of socioeconomic functioning was, therefore, Class II. A change in residence occurred during this interval when Mr. C. sold his home and bought a new one into which he moved. By definition, this did not influence the grade of socioeconomic functioning.

Mrs. L., a 53-year-old housewife, lived in a single home that was owned by her husband. She was functioning independently in her activities of daily living and received no agency or community financial support. Her level of socioeconomic functioning was, thus, Class II by the *Index of ISF*. She had an acute cerebral infarction, and 6 months later was in a nursing home where bathing and toileting functions required assistance by the nursing-home staff. She was not being dressed. Her husband paid all her expenses from his savings and salary. Her level of socioeconomic functioning was now Class IV. One year after the onset of her original illness, she was in the same incapacitated physical condition; however, she was now in a state hospital. Part of her financial support was assumed by the state; this established her level of socioeconomic functioning as Class V.

Mrs. C., a 95-year-old widow, lived with a wealthy daughter and son-in-law in a home owned by the latter. Mrs. C. was unemployed, and all financial needs were met by the children. She received some assistance in getting dressed, but bathed herself and went to the toilet independently. Her level of socioeconomic functioning was Class III according to the *Index of ISF*. Mrs. C. broke her hip in a fall and was hospitalized. After a prolonged convalescence, she showed no motivation to help herself in activities of daily living. Her daughter took her home and hired a full-time companion who assisted the patient in bathing, dressing, and toileting; thus, 6 months after fracture her level of socioeconomic functioning was Class IV. All financial needs continued to be met by the children, and there was no change in the level of socioeconomic functioning according to the *Index of ISF* during the next year. Mrs. C. died 21 months after the fracture in the home of her daughter.

Mrs. F., an 84-year-old widow, lived in an apartment that she did not own with a maid who acted as her companion, but who gave no personal assistance in activities of daily living. She was unemployed, yet financially completely independent with adequate savings and income from stocks and bonds. The fact that she did not own the residence in which she lived and the fact that she was unemployed established her level of socioeconomic functioning as Class III by the *Index of ISF*. She was discovered to have a breast cancer that later metastasized to the brain. She could no longer care for herself and was moved to a nursing home where she was a bed patient under the complete personal care of the nursing-home staff. Her level of socioeconomic functioning was now Class IV. There she remained until she died 18 months after the discovery of the cancer.

The index was used to study changes in patient status with regard to socioeconomic functioning following illness (Tables III and IV). A comparison of socioeconomic functioning or ISF ratings prior to illness with ratings 6 months later revealed 32 instances of deterioration among the 87 patients still alive at 6 months (Table III). In the higher ISF classes, namely, Classes I and II, 11 of the 36 patients (31 per cent) deteriorated to a lower class, while 21 of the 48 (44 per cent) in the lower classes deteriorated, namely, Classes III and IV. The 3 patients in Class V could not deteriorate because they were already in the lowest possible class. Data are not presented here as data from which to draw conclusions regarding the social impact of illness, but are presented to illustrate the possible use of the *Index of ISF* in studying this type of problem.

TABLE III. USE OF AN ORDERED INDEX TO COMPARE SOCIOECONOMIC FUNCTIONING PRIOR TO THE OCCURRENCE OF FRACTURE OF THE HIP AND 6 MONTHS LATER IN PATIENTS ADMITTED TO A HOSPITAL FOR PROLONGED ILLNESS

		Index of Independence in Socioeconomic Functioning PRIOR TO FRACTURE*						TOTAL
	CLASS	I	II	III	IV	V	OTHER	
Index of Independence in Socioeconomic Functioning 6 months after fracture	I	2						2
	II	1	23					24
	III		6	23				29
	IV		2	17	4			23
	V		1	1	3	3		8
	Other		1					1
	Dead		1	3	1	1		6
	Total	3	34	44	8	4	0	93

*See text and Table I for detailed definitions of classes comprising the *Index of Independence in Socioeconomic Functioning.*

TABLE IV. USE OF AN ORDERED INDEX TO COMPARE SOCIOECONOMIC FUNCTIONING 6 AND 12 MONTHS AFTER THE OCCURRENCE OF FRACTURE OF THE HIP IN PATIENTS ADMITTED TO A HOSPITAL FOR PROLONGED ILLNESS

		Index of Independence in Socioeconomic Functioning 6 MONTHS AFTER FRACTURE*						TOTAL
	CLASS	I	II	III	IV	V	OTHER	
Index of Independence in Socioeconomic Functioning 12 months after fracture*	I	2			1			3
	II		20	2	1		1	24
	III		2	22	1			25
	IV		1	2	17			20
	V			1		5		6
	Dead		1	1	3	3		8
	Total	2	24	28	23	8	1	86†

*See text and Table I for detailed definitions of classes comprising the *Index of Independence in Socioeconomic Functioning.*

†Six patients who had died during the first 6 months were not included in this table.

A comparison of ISF ratings 6 months after fracture with ratings 1 year after fracture revealed 6 instances of deterioration among the 78 patients still alive at 1 year and 6 instances of improvement (Table IV). The 6 instances of improvement occurred in patients who had shown deterioration during the first 6 months. There were 14 deaths during the year after fracture among the 92 consecutive patients who were evaluated at the 1 year level. The 14 deaths

comprised 6 per cent of those with higher ISF ratings prior to fracture, i.e., Classes I and II, and 21 per cent of those with lower ISF ratings, i.e., Classes III, IV, and V. Patients whose socioeconomic functioning deteriorated during the initial 6 months of illness had a greater mortality during the year following such deterioration than those without deterioration, the respective mortality percentages being 21 and 8 per cent.

As another example of the possible uses of the *Index of ISF* in studying problems relative to the social impact of illness, 6 month deteriorations in ISF ratings among 20 consecutive patients with recent, initial cerebral infarction were compared with 6 month ISF deteriorations among 93 consecutive patients with fracture of the hip. ISF deteriorations occurred in 80 per cent of the former group and in 38 per cent of the latter. In calculating these percentages, patients starting in the lowest class, V, were excluded since they were ineligible for further deterioration. Patients who died during the first 6 months were also excluded from the calculation. It was of methodologic interest to note that of the 38 evaluations of ISF status available in the preliminary analysis of patients with recent cerebral infarction, all 38 were classifiable by the *Index of ISF*.

A study of successive ISF ratings in individual patients disclosed that deteriorations in ISF ratings tended to occur only once during a succession of 3 or more observations. At least 3 successive ratings were available for analysis during life in each of 78 patients. Among the 78 patients, 66 had at least 4 successive ratings, 38 had at least 5 successive ratings, and 12 had 6 successive ratings. There were only 2 patients whose situtation deteriorated twice. In neither of the 2 exceptions was the deterioration progressive, since both were instances in which the patient deteriorated to a given level, then improved, and subsequently redeteriorated to the prior low level.

DISCUSSION

Speaking of a change in emphasis from facts of death to problems of disability in relation to chronic illness, Reed[6] states, "The rapidity with which this change has come about, together with the need for new yardsticks for measuring the effects of chronic disease, has made it difficult, even for students devoting themselves particularly to this problem, to get a real measure of the course of events." Recognizing the need for new yardsticks, members of the staff of The Benjamin Rose Hospital have been conducting a continuing program of investigation in this direction.[1] An ordered index was developed of the level of physical functioning of patients.[2] The index was termed the *Index of Independence in Activities of Daily Living* (*Index of ADL*) and its use as a longitudinal measure of disability was demonstrated.[3] The present report describes a newly evolved, ordered measure of certain aspects of socioeconomic functioning, termed the *Index of Independence in Socioeconomic Functioning* (*Index of ISF*). Included in the present report are examples of its possible uses as a longitudinal measure for studies of the socioeconomic impact of illness in various groups of patients and in various illnesses. Students of chronic disease such as Chinn,[7] Cosin,[8] and Rusk[9] have long indicated the need for tools by means of which relationships between social factors and illness could be demonstrated and studied further.

In the multivariate analysis used to evolve the *Index of ISF*, numerous items representing measures of resources, needs, and productivity were analyzed as described in the Methods section. The items that were finally combined into the index made it possible to rank patients in an ordered manner with respect to socioeconomic functioning. Included were defined grades of agency support, assistive nature of residence, residence ownership, and employment status. Although independently arrived at, the inclusion of the economic factors listed above is consistent with previous observations reported by Steiner.[4,5] On the basis of such experience as that derived from analysis of survey data on noninstitutionalized aged persons, Steiner noted that income differences among groups and individuals are largely explained by the presence and size of earnings, that differences in earnings reflect differences in labor force participation, that savings and assets have small impact on income except for a very few aged persons, and that the one important form of asset is typically the ownership of a house that is paid for.[4,5] Use of socioeconomic measures by others in relation to chronic disease includes measures such as an index of social ranking using education, occupation, and rental value or such as the Edwards' Occupational Categories.[10,11] These are aimed at classifying in terms of factors primarily indicative of productivity. The *Index of ISF*, evolved in the present study, aims at a combined, more over-all measure of productivity, needs, and resources. We consider that the method used to evolve the index is relatively unique as applied to the field of chronic disease, although methods involving factor analysis have been used previously to study economic patterns in the aged.[12] As described in the Methods section, the present derivation involves an ordered profiling of persons based on status defined in terms of graded definitions of individual variables. We have used this method in 2 instances to date, one to derive an over-all index of the level of activities of daily living, and the other to derive the current index of certain aspects of socioeconomic functioning. This method may have further application to the problem of establishing longitudinal yardsticks for the study of chronic disease with its complex multivariate relationships.

In the present report, illustrations were given of some possible applications of the *Index of ISF*. In study patients, comparison of ISF ratings prior to fracture of the hip with ISF ratings 6 months later revealed deterioration in 37 per cent of those still alive at 6 months. The proportion of deteriorations was greater in patients with lower ISF ratings prior to illness than in patients with higher ratings. A similar comparison of ratings 6 months after fracture with ratings 12 months after fracture showed deterioration in 8 per cent of those still alive at 12 months, as well as improvement in 8 per cent. The instances of improvement were observed in patients who had deteriorated during the first 6 months. Deterioration as measured by the ISF index was most likely to occur during the first interval of observation, was not progressive, and tended to persist. Of 14 patients who died during the first year after fracture, 2 had high ISF ratings prior to fracture while 12 had low prior ratings. Mortality, furthermore, appeared to be greater among patients who deteriorated in ISF rating than in those who did not. Another example of a use of the *Index of ISF* was as a means of comparing social impact of illness between different illnesses. A group of patients with

recent, initial episodes of cerebral infarction appeared to have more frequent ISF deteriorations than study patients who had fractured their hips. It is possible that the observed ratio of deteriorations between the 2 illnesses, 2:1, would be still larger if the numbers of patients available were large enough to permit analysis of age-corrected data.

Additional study of the *Index of ISF* is needed and is being conducted. The index was derived from data concerning aged patients with fracture of the hip and was applied to a group of patients with recent, initial cerebral infarction. Further tests of its applicability are required in other groups of healthy and sick aged persons in order to define its scope of usefulness, as well as its limitations. It may be, for example, that in certain geographic areas and in certain social and cultural settings, the *Index of ISF* as evolved here may have limited application. It may also be that limitations will be encountered in its use for certain purposes. We believe, on the other hand, that tools are not currently available for the critical study of socioeconomic functioning as it influences and is influenced by disease and disability. The effort described in this paper is an illustration of a way that the problem can be approached quantitatively and precisely. The *Index of ISF* is not described as a final solution to the problem; but, at this time, we consider the method to be one of great promise for use in studies of aged persons. Finally, we stress that we do not think of the *Index of ISF* as revealing social status, but our concept is that it is a measure of over-all socioeconomic *functioning*.

l. 13
961

SUMMARY

The staff of The Benjamin Rose Hospital, a hospital for prolonged illness in Cleveland, Ohio, has had special interest in developing much-needed tools, i.e., "yardsticks," for the study of the progression and regression of chronic disease. Methods of approaching this problem were designed initially in relation to a study of patients with fracture of the hip and extended subsequently to include patients with recent, initial episodes of cerebral infarction.

Previous reports defined and demonstrated the use of an over-all index of independence in activities of daily living, the *Index of ADL.*[2,3] In the present report, an *Index of Independence in Socioeconomic Functioning* (*Index of ISF*) was defined, and its derivation was described. The index is a composite of graded measures of certain factors representative of patient resources, needs, and productivity. More specifically, the factors in the index are agency support, assistive nature of residence, residence ownership, and employment status. The index was used to rank patients and to observe relative differences between various groups, as well as to observe changes in status from one time to another. The concept behind the development of an index of socioeconomic functioning was discussed, as well as its significance in approaching the problems of critical studies of relationships between socioeconomic functioning and disease or disability.

Examples of possible uses of the *Index of ISF* were presented. The following preliminary trends were revealed by these examples: more frequent 6 month deteriorations among patients with lower ISF ratings prior to fracture than among those with higher prior ratings; greater likelihood of deterioration during the

first 6 month interval after fracture than in subsequent 6 month intervals; lack of progression and persistence of deterioration when deterioration did occur; more frequent deaths during the first year after fracture among those with lower ISF ratings prior to fracture; more frequent fatalities observed within the year after the occurrence of ISF deterioration than observed in the absence of such deterioration; and more frequent ISF deterioration in the 6 month interval after recent, initial cerebral infarction than in the 6 months after fracture of the hip.

REFERENCES

1. The Staff of The Benjamin Rose Hospital: Multidisciplinary Study of Illness of Aged Persons. I. Methods and Preliminary Results, J. CHRON. DIS. **7**:332, 1958.
2. The Staff of The Benjamin Rose Hospital: Multidisciplinary Studies of Illness in Aged Persons. II. A New Classification of Functional Status in Activities of Daily Living, J. CHRON. DIS. **9**:55, 1959.
3. The Staff of The Benjamin Rose Hospital: Multidisciplinary Studies of Illness in Aged Persons. III. Prognostic Indices in Fracture of the Hip, J. CHRON. DIS. **11**:445, 1960.
4. Steiner, P. O., and Dorfman, R.: The Economic Status of the Aged, Berkeley and Los Angeles, Calif., 1957, University of California Press, pp. 38-152.
5. Donahue, W., and Tibbits, C.: The New Frontiers of Aging, Ann Arbor, Mich., 1957, University of Michigan Press, pp. 89-98.
6. Reed, L. J.: The Changing Role of Chronic Disease, J. CHRON, DIS. **1**:86, 1955.
7. Chinn, A. B.: Special Medical Problems in the Care of the Elderly Patient, Postgrad. Med. **24**:90, 1958.
8. Cosin, L. Z.: Physiotherapy in Geriatric Rehabilitation, Brit. J. Phys. Med. **18**:25, 1955.
9. Rusk, H. A., and Novey, J.: The Impact of Chronic Illness on Families, Marriage and Family Living **19**:193, 1957.
10. Ellis, J. M.: Socioeconomic Differentials in Mortality From Chronic Diseases, Social Problems **5**:30, 1957.
11. Graham, S.: Socioeconomic Status, Illness, and the Use of Medical Services, Milbank Mem. Fund Quart. **35**:58, 1957.
12. Millican, R. D.: Two Factor Analyses of Expenditures of the Aged, J. Gerontol. **14**:465, 1959.

Vol. 196

Book Reviews

THE TORCH. By Wilder Penfield, M.D. Boston, 1960, Little, Brown & Company. Pp. 370, indexed. Price $4.75.

This novel by a world-famous neurosurgeon and neurophyiologist deals with the life and times of Hippocrates. Dr. Penfield has apparently spent considerable time delving into the historical facts concerning the legendary if somewhat nebulous physician-teacher of Cos. Since we know little of the real Hippocrates, there is plenty of latitude available for a free-wheeling imaginative story. Unfortunately, Dr. Penfield is not exceptionally gifted as a novelist. As a result, this book rambles rather dully through a lackluster Asclepiad-finds-girl, Asclepiad-loses-girl, Asclepiad-gets-girl plot, unenlivened by perceptive character portrayal-in-depth. For physicians, there is a certain amount of interest to be derived from Penfield's descriptions of, and speculations about, the practice of medicine in ancient Greece, and the author has added an interesting set of historical notes and references, but the over-all effect is disappointing.

Louis Lasagna

CONTROLLED CLINICAL TRIALS. Papers delivered at the Conference convened by The Council for International Organizations of Medical Sciences. Prof. A. Bradford Hill, F.R.S., Chairman of the Conference. Springfield, Ill., 1960, Charles C Thomas. Pp. 177. Price $5.00.

This is a collection and reorganization of papers delivered at the Conference convened by the Council for International Organizations of Medical Sciences held in Geneva, 1959. This symposium was organized under the direction of Professor A. Bradford Hill, who, in addition to being the chairman, was one of the principal speakers. Its purpose was to discuss principles, organization, and scope of controlled clinical trials with drugs as well as ethical problems involved.

This small, easy-to-read book has been organized into nine parts. One part explains some of the techniques used in evaluating drugs in acute diseases. Other parts deal with chronic diseases such as pulmonary tuberculosis, rheumatoid arthritis, malignancy, and coronary thrombosis. Other topics discussed are: patients as their own controls, concurrent assessment of several treatments, and other designs in controlled clinical trials. Actual experiments have been included in almost every chapter to illustrate, in a practical way, the methodology and organization of each clinical trial. In some experiments the authors have indicated mistakes to be corrected or better techniques to be used in future trials.

Although only a few diseases are discussed, the investigative methods used in these will be applicable to others. Thus, the experimental clinician who has worked with new drugs in the past will find this book a useful manual for planning future experiments. The clinician who wishes to try his hand, for the first time, at a controlled clinical study will find the book equally useful as an introduction to this relatively new science.

Controlled Clinical Trials should remain a useful book for many years because it contains so many fundamental concepts for sound, reliable research. It would be, therefore, a good investment for those who are interested in clinical pharmacology.

John B. Murphy

CANCER AND ALLIED DISEASES OF INFANCY AND CHILDHOOD. Edited by Irving M. Ariel, M.D., and George T. Pack, M.D. Boston, 1960, Little, Brown & Company. Pp. 605, indexed. Price $22.50.

In this era, the publication of a new medical text has become an almost too commonplace event. At first glance, one might question the propriety of devoting an entire volume to a description of the various neoplastic diseases affecting children. However, a consideration of the causes of death in childhood reveals neoplasia as a common phenomenon which has been relatively neglected by pediatricians as a group.

Drs. Ariel and Pack have edited the contributions of more than a score of experts in the field and have drawn from their own extensive experience in formulating a general appraisal of the problem. The introductory chapters are concerned with the general care of children being treated for tumors and the basic principles of radiation therapy. The remainder of the book is divided according to the regional distribution of the various neoplasms, each section including a concise pathologic and clinical statement and a more detailed discourse on the appropriate operative, radiation, or drug therapy. Particular mention may be made of Dr. Melicow's chapter on the adrenal gland. The most frequent childhood malignancies, leukemia and lymphoma, are also discussed in detail.

This text clearly sets forth the problems of pediatric oncology for the first time. It is well illustrated and contains an extensive bibliography. It should prove quite useful to all those involved in the care of children with cancer and similar disorders and to those students who are seeking to acquaint themselves with this special field.

Albert H. Owens, Jr.

DRUGS AND BEHAVIOR. Edited by Leonard Uhr, Ph.D., and James G. Miller, M.D., Ph.D. New York, 1960, John Wiley and Sons, Inc. Pp. 676, indexed.

This is an ambitious volume, designed to provide "The methodological, chemical, biological, and clinical context for psychopharmacology" as well as "experimental procedures and results" in this growing field. The astonishing fact is that the book, despite its faults, has turned out to be a reasonably successful handbook for those workers in pharmacology and behavior who wish eclectic background in psychopharmacology.

The most useful parts of the book, in my opinion, are those which are most general in outlook, and least concentrated upon a single drug or procedure. For example, Nash's chapter on the design of experiments, Lorr's on rating scales, and Fleishman's compendium of psychomotor tests should be helpful to many people. In contrast, I doubt that the editors could justify the inclusion of a one-and-a-half-page "chapter" on the effects of oxanamide on a simulated driving task.

There is a fair amount of material in the book with which one might take issue. For example, Waggoner's statement that "psychotherapy is still . . . the most important and best method" of treating mental illness certainly needs strong qualification, at least in regard to schizophrenia and certain affective disorders. The editors' concept of constructing "nationwide" norms which would then eliminate the need for simultaneous, contemporaneous controls is a naive and dangerous one. Some of the experimental data, such as those of Olds and Travis, approach dose-response problems in a remarkably unpharmacologic way. Eysenck's "deductions" involving "stimulant" and "depressant" drugs deal with pharmacologic agents in a conceptual manner which flamboyantly disregards the rich variety of complex changes produced by any drug. Porteus' pronouncements on placebos will dismay some readers. And so on—nevertheless, there is much in this book of interest and utility and psychopharmacologists should at least browse through its pages.

Louis Lasagna

SURGICAL DISEASES OF THE PANCREAS. By John M. Howard, M.D., F.A.C.S., and George L. Jordan, Jr., M.D., M.S., F.A.C.S. Philadelphia and Montreal, 1960, J. B. Lippincott Company. Pp. 607, indexed. Price $20.00.

As indicated in the Foreword, the book was written with the purpose of concentrating in one volume the essential knowledge concerning the diseases of the pancreas, excluding diabetes

mellitus. For the most part, the authors have succeeded admirably. The contents provide a valuable and detailed reference work for all physicians interested in pancreatic disease, and the material is not limited to the domain of the surgeon as the title may suggest.

The volume is printed on excellent-quality paper with large type and outstanding illustrations. There are numerous tables to document the text, and an extensive bibliography accompanies each chapter, one chapter alone requiring 364 references. This same chapter on "Hyperinsulinism and Islet Cell Tumors of the Pancreas" contains a series of tables covering 82 pages listing all of the tumors reported in the literature associated with the pancreatic islet cells or hypoglycemia. This detailed information may possibly increase the value of the book but most certainly increases its price.

The authors elected to omit the field of animal experimentation and have reported primarily clinical experiences. They recognize that data accumulated on the animal may not be comparable to observations in the human. Nonetheless, this omission will be regretted by readers interested in all aspects of the subject, particularly those engaged in research.

A rather startling deficiency is the handling of the problem of pancreatic biopsy. This is a common problem of extreme importance to the clinical surgeon and begs for detailed analysis and documentation. Lamentably, the subject is covered in a single paragraph in the section on "Pancreatic Resection" and is not listed in the index.

For the reader who desires a reference work, the contents will be most satisfactory. When read from cover to cover, however, there is considerable repetition, an almost inevitable phenomenon when multiple contributors are writing about a single organ. Dr. Howard covers the sections dealing with various phases of pancreatitis, a field in which he has had considerable personal experience and in which he has published numerous original papers. The association of acute pancreatitis with either alcoholism or gall stones is accentuated repeatedly, and the clinical differentiation of the two forms is convincingly documented. Many gastroenterologists will consider this approach an oversimplification. Dr. Jordan covers cysts, tumors, and surgery of the pancreas, while a number of other distinguished contributors cover the history, anatomy, embryology, anomalies, and physiology of the pancreas.

This will be a difficult book to revise because of the numerous references to the current literature. Perhaps the authors in the future will summarize and consolidate more thoroughly the varying views and results instead of enumerating each paper individually. Such editing, if critically done, would add to the readability of the text without detracting from its accuracy or completeness.

Despite the deficiencies outlined, this is an excellent treatise which every good medical library should obtain, if only for its reference value. The abdominal surgeon and the internist can benefit from a careful reading of this book but must be prepared for rather heavy going in some areas, and insufficient coverage in others.

Howard Patt

Vol. 13
1961

Announcements

The Department of Physical Medicine and Rehabilitation of the New York Medical College–Metropolitan Medical Center announces a one-week course in ELECTRODIAGNOSIS AND ELECTROMYOGRAPHY to be offered June 5 to 9, 1961.

This course for physicians is planned to provide detailed instruction in the principles, techniques, and instrumentation involved in electrodiagnosis and electromyography. The course will consist of lectures, seminars, clinical demonstrations, and laboratory sessions. The teaching staff includes members of the faculty of the medical College and noted guest lecturers.

The curriculum will include:

Fundamentals of Nerve and Muscle Physiology
Electrical Activity of Nerve and Muscle
Electronics
Instrumentation
Electrodiagnosis and Electromyography: Applications
Clinical Laboratory
Special Techniques
Clinical and Research Implications

For further information, please contact Raymond C. Lerner, M.S.S.W., Coordinator, Postgraduate Education, Department of Physical Medicine and Rehabilitation, New York Medical College, 1 East 105th St., New York 29, N. Y.

The University of Colorado School of Medicine announces the COCHEMS COMPETITION, funds for which were provided in the will of the late Mrs. Jane Nugent Cochems. A prize of $2,500 will be awarded to the author of the best paper on the subject of "The Diagnosis, Etiology, and Treatment of Thrombophlebitis." The competition is open to all physicians, and entries must be received, in triplicate, on or before Oct. 1, 1961.

The Colorado National Bank of Denver, trustee under the will of Jane Nugent Cochems, has requested the Dean of the University of Colorado School of Medicine to conduct the competition. The judges, appointed by him, are Dr. Michael E. De Bakey, Professor and Head of the Department of Surgery, Baylor University College of Medicine, and Dr. Sol Sherry, Professor of Medicine, Washington University School of Medicine.

Papers submitted in the competition may not be published until after the winner of the competition has been announced. At that time, the winning paper and all others may be published at the discretion of individual authors. It should be noted, however, that those involved in conducting the competition will not assume any responsibility for submitting manuscripts for publication nor for any costs incident thereto. The winning paper, if published, must carry the designation, "Awarded the Jane Nugent Cochems Prize."

Further questions regarding the competition and all manuscripts should be directed to Dr. Robert J. Glaser, Vice-President for Medical Affairs and Dean of the University of Colorado School of Medicine, University of Colorado Medical Center, 4200 East Ninth Avenue, Denver 20, Colo.